# DATA DETECTION IN MASSIVE MU-MIMO SYSTEMS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Charles Jeon

May 2019

DATA DETECTION IN MASSIVE MU-MIMO SYSTEMS

Charles Jeon, Ph.D.

Cornell University 2019

Massive multi-user (MU) multiple-input multiple-output (MIMO) will be a core technology in next-generation wireless systems. By equipping the infrastructure base-stations (BSs) with hundreds of antenna elements and serving tens of user equipments (UEs) in the same time-frequency resource, massive MU-MIMO enables orders-of-magnitude higher spectral efficiency than existing wireless systems. The presence of large number of antenna elements at the BS, however, causes significant implementation challenges. In particular, optimal data detection at the BS that maximizes the spectral efficiency (i.e., the number of bits that can be transmitted reliably over a given bandwidth) entails prohibitive complexity. As a result, the majority of existing data detection algorithms for massive MU-MIMO and corresponding hardware designs are sub-optimal, thereby sacrificing spectral efficiency.

In this thesis, we will provide a positive answer to the question "Is optimal data detection in massive MU-MIMO systems feasible?" by considering a multi-disciplinary research approach that spans theory, algorithm development, and application-specific integrated circuit (ASIC) design. Concretely, we will propose a range of solutions on theory, algorithm, and hardware level that enable optimal data detection in practice. In addition, we will present new methods that reduce the complexity of channel-matrix preprocessing, as well as novel architectures and algorithms that enable parallel processing of the most critical tasks in massive MU-MIMO BSs. In order to demonstrate the effectiveness of all our solutions in practically-relevant communication scenarios, we will support our findings via theoretical results, numerical simulations, and ASIC implementations.

## BIOGRAPHICAL SKETCH

Charles Jeon obtained his BS with a dual major in Electrical Engineering and Mathematics graduating *magna cum laude*, and MS in Electrical Engineering from the University of Pennsylvania in 2013. During the same year, he also received a BS in Economics with concentrations in Finance and Statistics from The Wharton School. Since finishing his undergraduate studies, Charles has been pursuing a PhD at Cornell University in School of Electrical and Computer Engineering, where he worked on novel data detection methods and digital circuit implementations for next-generation wireless systems. During his PhD work, Charles has held internships at Xilinx, FutureWei Technologies, and Intel Labs; prior coming to Cornell, he has held internships at MIT Lincoln Laboratory, Deutsche Bank, and Hanwha Securities. Charles was the recipient of John McMullen Dean's Fellowship, ECE Hewlett Packard Fellowship, and Cornell ECE Outstanding PhD TA award.

*To my family for their unconditional love and support.*

## ACKNOWLEDGEMENTS

This dissertation was a culmination of many years of work across multiple disciplines–ranging from mathematical proofs, computer simulations, and circuit-level descriptions. Learning and performing research across these disciplines would not have been possible without the help of Professor Christoph Studer. I would like to thank him for taking me as his first student without any knowledge in digital circuits. Coming from a half-finance and half-engineering background, I asked Christoph many questions on how wireless systems work, and how even the most simplistic arithmetic operations were implemented in hardware. Christoph was always patient with me, and never hesitated to describe other methods for describing the same functionality. Throughout my journey in graduate school, Christoph has instilled me the importance of cross-layer optimization and the importance of building devices that actually work. His passion for signal processing was present anywhere and anytime, whether it be when eating chicken nuggets at McDonalds on a Sunday lunchtime, or drinking Sangrias at CollegeTown Bagels on Fridays past midnight. I will never forget the research discussions we had, and your teachings will resonate throughput my career. Thanks again Christoph!

My graduate school career would not have been successful without the teachings of my undergraduate academic advisors at the University of Pennsylvania. Professor Santosh Venkatesh has taught me the importance of intuitive thinking and the beauty of mathematical analysis. The Fourier analysis and probability courses were difficult, but I would not be here without solving his elegant mathematical puzzles. I would like to thank Professor Saleem Kassam for teaching me both analog to digital communications. Whenever I dropped by, he would provide me with wisdom not only about signal processing, but also about life in

general. I will never forget his teachings for using mathematics as a tool. I would also like to thank Professor Jan Van der Spiegel for always giving me research feedback when I visited him. Everytime I visited his office after graduating Penn, he provided me with directions on how I should think as a circuit designer for my next research in signal processing.

I am indebted to my colleagues and friends for not only teaching me about research, but also providing me wisdom on life directions after graduate school. Ramina Ghods, it was a pleasure to discuss research problems with you; you helped me to actually sit and think about problems that I would have glanced over if I did not discuss it with you. Nirmal Shende, it was unfortunate that we did not have a chance to write a paper together in graduate school. I had a great time with you at ISIT in Barcelona, whether it be practicing together or attending each other's conference talks. Shih-Hao Tseng, we had a great time touring each other's hometown in Taipei and Seoul. Next time, I will take you to the best Taiwanese restaurant in Korea! Ecenur Ustun, you always greeted me with a huge smile anytime. Even though I would be having a bad day, you would always come by and say "Hi$\sim$" to cheer me up. I wish you carry the positive vibe anywhere you go, and hope that you continue to listen to KPop. Ryan O'Hern, you have introduced me to serious lifting, and helped me become stronger both physically and mentally. I will never forget the knowledge that you have taught me over coffee (which includes lifting supplements), and to become a better man. Ritchie Zhao, thanks for being a great housemate, and more importantly, a great friend. I initially did not understand why kitchen top spaces are important, but now I believe it because you introduced me how to cook systematically. I will never forget our great times making (spicy) Korean food at Zhao's kitchen. Khalid Al-Hawaj, we had a blast during graduate school. I would like to thank

you for providing me with wisdom on both work and non-work related things, and how to have a colorful and fulfilling life. I will never forget the numerous discussions we had, whether it be while eating Oreo McFlurries at a parking lot at 11:59pm, or drinking coffee until dawn outside during cold Ithaca winters. I would like to thank again all my friends in Computer Systems Laboratory at Cornell. I will never forget all the memories that I had, and the countless number of cups of coffee I shared with all of you.

Most importantly, I would like to thank my family for their unconditional love and support. Their sacrifices for my education is the reason why I was able to study and learn in the US. Thanks for staying with me in both good and bad times. I cannot express my gratefulness to my family in words.

## TABLE OF CONTENTS


Biographical Sketch . . . iii
Dedication . . . iv
Acknowledgements . . . v
Table of Contents . . . viii
List of Tables . . . xiii
List of Figures . . . xiv

**1 Introduction** **1**
1.1 Abstract . . . 1
1.2 Contributions in Massive MU-MIMO . . . 4
1.2.1 Individually-Optimal Data Detection . . . 4
1.2.2 Mismatched Data Detection . . . 7
1.2.3 Decentralized Data Detection . . . 10
1.2.4 Efficient Preprocessing . . . 13
1.2.5 ASIC Design of Optimal Data Detection . . . 15
1.2.6 ASIC Design of a Nonparametric Equalizer . . . 17
1.3 Notation . . . 19
1.4 Full Publication List . . . 20

**2 Massive MU-MIMO: Basics and State-of-the-Art** **23**
2.1 Motivation for Massive MU-MIMO . . . 23
2.2 MIMO and Massive MU-MIMO Wireless Technology . . . 24
2.2.1 Massive MU-MIMO . . . 26
2.3 Introduction to Wireless Systems . . . 29
2.4 System Model . . . 31
2.5 The Basics of MIMO Data Detection . . . 33
2.5.1 Optimal Data Detection . . . 33
2.5.2 Linear Data Detection . . . 36

**3 Individually-Optimal Data Detection in Massive MU-MIMO** **42**
3.1 Introduction . . . 42
3.1.1 Application Examples . . . 43
3.1.2 Relevant Prior Art . . . 45
3.2 Complex Bayesian Approximate Message Passing (cB-AMP) . . . 46
3.2.1 System Model and Assumptions . . . 47
3.2.2 cSE: Complex State Evolution (with Mismatch) . . . 51
3.3 LAMA: Large MIMO Approximate Message Passing . . . 54
3.3.1 Large MIMO and Optimal Data Detection . . . 54
3.3.2 Derivation of the LAMA Algorithm . . . 56
3.3.3 LAMA Decouples Large-MIMO Systems . . . 58
3.3.4 LAMA and MF Data Detection . . . 62
3.3.5 LAMA in the Massive MU-MIMO Limit . . . 63

3.4 Optimality of LAMA . . . 64
3.4.1 Existing Results of IO and Multiuser Detection . . . 65
3.4.2 Fixed Points of LAMA . . . 67
3.4.3 When Does LAMA Achieve IO Data Detector Performance? 68
3.4.4 Optimality Conditions for LAMA with No Noise . . . 71
3.4.5 Optimality Conditions for LAMA With Noise . . . 73
3.4.6 Decomposing Complex-Valued Systems . . . 81
3.4.7 ERT, MRT, and Critical Noise Levels . . . 84
3.5 Numerical Asymptotic Results . . . 85
3.5.1 Achievable Rates and Error-Rate Performance . . . 86
3.5.2 Performance and Complexity Trade-offs . . . 88
3.5.3 Performance in Finite-Dimensional Systems . . . 91
3.5.4 Extension to General Channel Matrices $\mathbf{H}$ . . . 92
3.5.5 Simulation Results . . . 93
3.6 LAMA and Prior Art . . . 97
3.6.1 BPSK Signaling in Randomly Spread CDMA systems . . . 97
3.6.2 Recovery of Antipodal Solutions via Convex Optimization 100
3.7 Conclusions . . . 101

**4 Mismatched Data Detection in Massive MU-MIMO 102**
4.1 Introduction . . . 102
4.1.1 Why Should One Use a Mismatched Prior? . . . 103
4.1.2 Relevant Prior Art . . . 104
4.2 Mismatched Complex Bayesian AMP . . . 105
4.2.1 The mcB-AMP Algorithm . . . 105
4.2.2 Mismatched State Evolution Framework . . . 106
4.2.3 Optimal Tuning of the Variance Parameter $\tau$ . . . 109
4.2.4 Fixed-point Analysis . . . 110
4.3 Mismatched Data Detection with Optimal Tuning: General Case 111
4.3.1 Optimally-Tuned Data Detection with a Gaussian Prior . . 111
4.3.2 Suboptimal Data Detection with a Gaussian Prior . . . 113
4.4 Mismatched Data Detection with Optimal Tuning: QAM/PAM . 114
4.4.1 Optimally-Tuned Data Detection with a Hypercube Prior . 115
4.4.2 Suboptimal Data Detection with a Hypercube Prior . . . 116
4.4.3 Optimally-Tuned Data Detection with Gray Coding . . . 119
4.4.4 Suboptimal Data Detection with Gray Coding . . . 123
4.5 Numerical Results . . . 124
4.6 Conclusions . . . 127

**5 Decentralized Data Detection in Massive MU-MIMO 128**
5.1 Introduction . . . 128
5.1.1 Decentralized Baseband Processing . . . 129
5.1.2 Relevant Prior Art . . . 130
5.2 Decentralized Equalization Architectures . . . 131

5.2.1 Uplink System Model and Equalization . . . 133
5.2.2 Basics of Decentralized Equalization . . . 134
5.2.3 Partially Decentralized (PD) Equalization Architecture . . 134
5.2.4 Fully Decentralized (FD) Equalization Architecture . . . . 135
5.2.5 Signal-to-Interference-and-Noise-Ratio (SINR) Analysis . 136
5.3 Partially Decentralized (PD) Equalization . . . 137
5.3.1 Linear Equalization Algorithms for the PD Architecture . 137
5.3.2 LAMA for the PD Architecture . . . 141
5.4 Fully Decentralized (FD) Equalization . . . 142
5.4.1 Optimal Fusion for the FD Architecture . . . 142
5.4.2 SINR Analysis of Optimal Fusion in the FD Architecture . 144
5.4.3 Antenna Partitioning Strategies for Linear Equalizers . . . 145
5.5 Numerical Results . . . 147
5.5.1 Achievable Rate Analysis . . . 148
5.5.2 Asymptotic vs. Finite-Dimensional Systems . . . 153
5.5.3 Coded Error-rate Performance in Realistic Systems . . . . 153
5.6 Conclusions . . . 155

**6 Efficient Preprocessing For Massive MU-MIMO 156**
6.1 Introduction . . . 156
6.1.1 Interpolation-Based Matrix Computations . . . 157
6.1.2 Relevant Prior Art . . . 158
6.2 Prerequisites . . . 159
6.2.1 System Model . . . 159
6.2.2 Wideband Channel Model . . . 160
6.2.3 Linear and Non-linear Data Detection and Precoding . . . 162
6.3 Interpolation-based Gram Matrix Computation . . . 163
6.3.1 Exact Gram-Matrix Interpolation . . . 163
6.3.2 Approximate Gram-Matrix Interpolation . . . 165
6.4 Approximation Error Analysis . . . 167
6.4.1 Mean-Square-Error of Approximate Interpolation . . . 167
6.4.2 Mean-Square-Error of 0th Order Interpolation . . . 168
6.4.3 Mean-Square-Error of 1st Order Interpolation . . . 170
6.5 Complexity Analysis . . . 172
6.5.1 Brute-Force Computation . . . 173
6.5.2 Exact Interpolation . . . 173
6.5.3 0th Order Interpolation . . . 174
6.5.4 1st Order Interpolation . . . 175
6.6 Numerical Results . . . 175
6.6.1 Complexity Comparison . . . 176
6.6.2 MSE of Approximate Interpolation . . . 177
6.6.3 Error-rate Performance . . . 178
6.6.4 Performance and Complexity Trade-offs . . . 182
6.7 Conclusions . . . 183

**7 ASIC Design of Optimal Data Detection for Massive MU-MIMO 185**
7.1 Introduction . . . 185
7.1.1 Relevant Prior Art . . . 186
7.2 Massive MIMO Detection . . . 188
7.2.1 System Model . . . 188
7.2.2 Iterative MIMO Detection and Decoding . . . 188
7.3 Robust Algorithm Optimizations for LAMA . . . 189
7.3.1 Algorithmic Improvements For General Channels . . . 191
7.3.2 Efficient Posterior Mean and Variance Computation . . . 192
7.3.3 Preprocessing . . . 196
7.3.4 Architecture Overview . . . 198
7.3.5 Posterior Function and Onsager Computation . . . 202
7.3.6 Matrix-vector Multiplier Unit . . . 205
7.3.7 SINR Computation: Newton-Raphson Based Reciprocal . 206
7.3.8 LLR Computation Unit . . . 208
7.4 Implementation Results . . . 208
7.4.1 Post-placement Layout Results . . . 209
7.4.2 Fixed-point Error Rate Performance . . . 211
7.4.3 Comparisons to Other Detector Implementations . . . 216
7.5 Conclusions . . . 218

**8 ASIC Design of a Nonparametric Equalizer for Massive MU-MIMO 220**
8.1 Introduction . . . 220
8.1.1 The Case for Nonparametric Equalization . . . 220
8.2 A Primer on L-MMSE Equalization . . . 221
8.2.1 System Model . . . 222
8.2.2 Basics of L-MMSE Equalization . . . 222
8.2.3 L-MMSE Equalization via mcB-AMP . . . 223
8.3 Nonparametric Equalizer (NOPE) . . . 224
8.3.1 The NOPE Algorithm . . . 225
8.3.2 Robust Version of NOPE . . . 225
8.3.3 Numerical Results . . . 228
8.4 VLSI Architecture and Synthesis Results . . . 229
8.4.1 Architecture Overview . . . 230
8.4.2 Architecture Details . . . 231
8.4.3 Implementation Results and Conclusions . . . 235

**9 Conclusions and Outlook 236**

**A Proofs 239**
A.1 Proofs for Chapter 3 . . . 239
A.1.1 Proof of [1, Lem. 5.56] for complex-valued systems . . . 239
A.1.2 Derivation of Algorithm 1 . . . 240
A.1.3 Proof of Lemma 1 . . . 245

A.1.4 Intuitive derivation of Theorem 2 . . . 245
A.1.5 Proof of Lemma 4 . . . 247
A.1.6 Proof of Lemma 5 . . . 248
A.1.7 Proof of Lemma 6 . . . 249
A.1.8 Proof of Lemma 9 . . . 250
A.1.9 Proof of Theorem 10 . . . 251
A.1.10 Proof of Lemma 11 . . . 251
A.1.11 Proof of Lemma 12 . . . 251
A.1.12 Proof of Theorem 13 . . . 252
A.1.13 Proof of Lemma 14 . . . 254
A.1.14 Proof of Lemma 15 . . . 255
A.2 Proofs for Chapter 4 . . . 256
A.2.1 Proof of Lemma 17 . . . 256
A.2.2 Proof of Lemma 20 . . . 256
A.2.3 Proof of Lemma 21 . . . 256
A.2.4 Proof of Lemma 23 . . . 258
A.2.5 Proof of Lemma 25 . . . 258
A.3 Proofs for Chapter 5 . . . 261
A.3.1 Derivation of Algorithm 4 . . . 261
A.3.2 Proof of Lemma 28 . . . 261
A.3.3 Proof of Lemma 29 . . . 262
A.3.4 Proof of Theorem 30 . . . 262
A.3.5 Proof of Lemma 31 . . . 263
A.3.6 Proof of Lemma 32 . . . 264
A.3.7 Proof of Lemma 33 . . . 264
A.3.8 Proof of Lemma 34 . . . 264
A.3.9 Proof of Lemma 35 . . . 265
A.4 Proof of Chapter 6 . . . 266
A.4.1 Proof of Lemma 36 . . . 266
A.4.2 Proof of Lemma 37 . . . 266
A.4.3 Proof of Theorem 38 . . . 267
A.4.4 Proof of Corollary 40 . . . 269
A.4.5 Proof of Theorem 41 . . . 270
A.4.6 Proof of Corollary 42 . . . 270
A.5 Proofs of Chapter 7 . . . 272
A.5.1 Proof of Theorem 43 . . . 272

**Bibliography** **274**

## LIST OF TABLES


3.1 Summary of (Sub-)Optimality Regimes of LAMA . . . . . 73
3.2 ERTs $\beta_{\mathcal{O}}^{\max}$, MRTs $\beta_{\mathcal{O}}^{\min}$ and the critical noise levels $N_0^{\min}(\beta_{\mathcal{O}}^{\min})$ and $N_0^{\max}(\beta_{\mathcal{O}}^{\max})$ for LAMA with common PSK, PAM, and QAM constellations . . . . . 83

7.1 Detailed area and power breakdown of the post-placement layout of LAMA . . . . . 210
7.2 Post-layout results of LAMA and comparison to existing massive MU-MIMO data detectors . . . . . 219

8.1 Synthesis results of NOPE for a 64 BS antenna, 16 UE system and comparison to existing massive MU-MIMO data detectors. . . . 234

## LIST OF FIGURES


1.1 Contributions overview . . . . . 4
1.2 Contributions of Chapter 3 . . . . . 5
1.3 Contributions of Chapter 4 . . . . . 8
1.4 Contributions of Chapter 5 . . . . . 11
1.5 Contributions of Chapter 6 . . . . . 15
1.6 Contributions of Chapter 7 . . . . . 16
1.7 Contributions of Chapter 8 . . . . . 18

2.1 Measured global monthly cellular traffic data/voice . . . . . 24
2.2 Example of a MIMO wireless device: Samsung Galaxy S8 . . . . 25
2.3 Trends and data rates of modern wired and wireless technology 25
2.4 Illustration of conventional MIMO and massive MIMO . . . . . 27
2.5 Illustration of single-user and multi-user massive MIMO . . . . . 29
2.6 Illustration of fading in a wireless system . . . . . 30
2.7 LoS and NLoS propagation effects . . . . . 30
2.8 Input-output system model for the MU-MIMO uplink . . . . . 32

3.1 LAMA decouples large-MIMO systems . . . . . 58
3.2 MIMO system with LAMA and its equivalent decoupled system 61
3.3 Function (3.34) for $\beta < \beta_{\text{QPSK}}^{\min}$ . . . . . 77
3.4 Function (3.34) for $\beta \in [\beta_{\text{QPSK}}^{\min}, \beta_{\text{QPSK}}^{\max}]$ . . . . . 78
3.5 Function (3.34) for $\beta > \beta_{\text{QPSK}}^{\max}$ . . . . . 79
3.6 Function (3.34) for 16-PSK and $\beta = 1.73 \in [\beta_{\text{16-PSK}}^{\min}, \beta_{\text{16-PSK}}^{\max}]$ . . . . 80
3.7 Achievable rate of LAMA for different system ratios with QPSK 87
3.8 Symbol error rate of LAMA for different system ratios with QPSK 88
3.9 Performance/complexity trade-offs of LAMA for different system ratios with QPSK . . . . . 89
3.10 Symbol error rate (SER) of LAMA in the large-system limit compared to the optimal SER and the SER of an AWGN channel . . . 94
3.11 Performance/complexity trade-off of LAMA in the analytical large-system limit and simulated finite dimension . . . . . 95
3.12 Symbol error rate of LAMA for a $128 \times 64$ system with 16-QAM 96
3.13 Symbol error rate of LAMA for a $128 \times 128$ system with 16-QAM 97
3.14 Symbol error rate of LAMA for a $128 \times 128$ system with QPSK and channel estimation errors . . . . . 98

4.1 Symbol error rate for M-LAMA algorithms for $128 \times 64$ MIMO system with QPSK and 10 algorithm iterations . . . . . 125
4.2 Symbol error rate for M-LAMA algorithms for $128 \times 64$ MIMO system with 16-QAM and 5 algorithm iterations . . . . . 126

5.1 Partially and fully decentralized feedforward equalization architectures for the massive MU-MIMO uplink . . . . . 132

5.2 Achievable rate analysis of feedforward decentralized baseband processing for QPSK: fixed target rate of $R = 1.99$ . . . . . 148
5.3 Achievable rate analysis of feedforward decentralized baseband processing for QPSK: fixed SNR loss of 1 dB . . . . . 149
5.4 Achievable rate analysis of feedforward decentralized baseband processing for 16-QAM: fixed target rate of $R = 3$ . . . . . 150
5.5 Achievable rate analysis of feedforward decentralized baseband processing for 16-QAM: fixed SNR loss 1 dB . . . . . 151
5.6 SER of decentralized equalization in the large-system limit and $256 \times 16$ Rayleigh-fading massive MU-MIMO system . . . . . 152
5.7 Packet error-rate of an LTE-like $64 \times 16$ massive MU-MIMO-OFDM system in WINNER II channel model . . . . . 154

6.1 Illustration of 0th and 1st order interpolation across subcarriers . 166
6.2 Comparison of Gram matrix computational complexity . . . . . 177
6.3 MSE of 0th and 1st order interpolation . . . . . 179
6.4 Bit error rate for $128 \times 8$ massive MU-MIMO-OFDM system for 16-QAM with brute-force and approximate Gram matrix interpolation for i.i.d. Rayleigh fading with perfect CSI . . . . . 180
6.5 Bit error rate for $128 \times 8$ massive MU-MIMO-OFDM system for 16-QAM with brute-force, exact, and approximate Gram matrix interpolation for i.i.d. Rayleigh fading with channel estimation . 181
6.6 Bit error rate for $128 \times 8$ massive MU-MIMO-OFDM system for 16-QAM with brute-force, exact, and approximate Gram matrix interpolation for QuaDRiGa channel with channel estimation . . 182
6.7 Trade-off between SNR performance and computational complexity for 0th and 1st order interpolation in an i.i.d. Rayleigh fading channel with channel estimation errors . . . . . 183

7.1 Coarse-grained pipeline architecture for LAMA . . . . . 198
7.2 Top-level diagram of LAMA for Algorithm 5 . . . . . 200
7.3 Log-likelihood Ratio unit for $b$th bit in posterior function unit . . 202
7.4 Post-placement layout of the LAMA chip . . . . . 209
7.5 $256 \times 32$ MU-MIMO-OFDM system with i.i.d. Rayleigh fading . 212
7.6 $256 \times 32$ MU-MIMO-OFDM system with Urban Micro channel . 213
7.7 $32 \times 32$ MU-MIMO-OFDM system with i.i.d. Rayleigh fading . . 214
7.8 $32 \times 32$ MU-MIMO-OFDM system with Urban Micro channel . . 215

8.1 Bit error rate of NOPE algorithm in a $64 \times 16$ massive MU-MIMO system with i.i.d. Rayleigh fading . . . . . 228
8.2 Top level block diagram of NOPE algorithm . . . . . 229
8.3 Illustration of the matrix-vector unit (MVU) that computes computes $\mathbf{H}\hat{\mathbf{s}}$ and $\mathbf{H}^{\mathrm{H}}\mathbf{r}$ in a $3 \times 3$ system via Cannon's algorithm . . . 230
8.4 Post-placement layout of the NOPE chip . . . . . 233

CHAPTER 1

# INTRODUCTION

## 1.1 Abstract

Wireless technology has transformed the way people connect, work, and play. For example, high-definition video calls can be done over a smartphone nearly anytime and anywhere—gamers can enjoy real-time high-fidelity worldwide gaming in a portable setup, whether it is on a smartphone or a laptop. The transformative impact of wireless technology on society is expected to continue with the introduction of novel applications, including virtual and augmented reality, autonomous driving, industry 4.0, and the Internet of things (IoT).

In order to enable these applications while supporting the exponential growth in the number of connected devices, without sacrificing data rates, energy efficiency, and quality-of-service, novel wireless technologies are necessary. Massive multi-user (MU) multiple-input multiple-output (MIMO) is a promising technology as it enables the communication between an infrastructure base-station (BS) and tens to hundreds of user equipments (UEs) in the same time-frequency resource. The idea of this technology is to equip the BS with an antenna array consisting of hundreds to thousands of antenna elements, which enables fine-grained beam-forming (i.e., focusing the transmitted energy towards the UEs), mitigates MU interference, and improves energy efficiency. Although massive MU-MIMO is a candidate technology for fifth-generation (5G) wireless systems, the presence of a large number of antenna elements and the need to sustain high data rates results in significant implementation challenges. One critical challenge is the computational complexity of optimal data detection, which in-

creases exponentially in the number of transmitting UEs. As a result, the majority data detection algorithms proposed for massive MU-MIMO and corresponding hardware design are sub-optimal in terms of the spectral efficiency.

Sub-optimal linear data detection algorithms are known to achieve optimal performance in massive MU-MIMO limit, i.e., where one fixes the number of UEs and increases the number of BS antennas to infinity. However, in realistic, finite-dimensional systems where the number of UEs is comparable to the number of BS antennas, linear methods perform only poorly. In fact, only optimal data detection maximizes the spectral efficiency and is able to fully exploit the capabilities of the large antenna array at the BS. This maximization in spectral efficiency is crucial for BS operators as more UEs can be served simultaneously by the same equipment, which leads to reduced operating cost per UE and an increase in revenue. In addition, optimal data detection enables one to operate at lower signal-to-noise-ratio (SNR) levels while maintaining highest spectral efficiency, which improves throughput, coverage, range, and quality-of-service. Moreover, UEs can communicate at lower transmit power, which has the potential to increase battery life—a critical aspect of devices for the IoT. Despite all these advantages, the design of optimal data detectors for massive MU-MIMO was, until now, an elusive goal.

This thesis boldly addresses the question: "Is optimal data detection in massive MU-MIMO systems feasible?" This thesis provides a positive answer to this question by considering a multidisciplinary research approach that spans theory, algorithm development, and hardware design. Such a holistic approach is necessary as the problem of optimal data detection is multifaceted: (i) optimal data detection entails prohibitive complexity, (ii) robust yet scalable algorithms

that perform well in a range of application scenarios are required, and (iii) corresponding algorithm solutions must be hardware friendly in order to enable an efficient integration in digital integrated circuits. Unfortunately, simply optimizing and tuning existing algorithm solutions and their corresponding hardware implementations is not a viable solution. As a result, this thesis approaches this problem by proposing a completely novel solution via the joint consideration of theory, algorithms, and hardware design.

This thesis develops optimal data detection algorithms for massive MU-MIMO and provides corresponding theoretical performance guarantees. In order to assess the real-world benefits of the proposed solutions (in terms of silicon area, throughput, and power consumption), prototype hardware designs in application-specific integrated circuits (ASICs) will be presented. In addition, the trade-offs between optimality and complexity are investigated by systematically analyzing suboptimal data detection methods that are able to achieve near-optimal performance while reducing hardware complexity. We also present new methods that reduce the preprocessing complexity of data detection and novel architectures and algorithms that enable parallel processing on multiple computing fabrics, which facilitates practical BS designs. For all of the proposed methods, a thorough theoretical analysis in combination with extensive numerical simulations and comparisons with existing algorithms for a range of practically-relevant communication scenarios are provided.

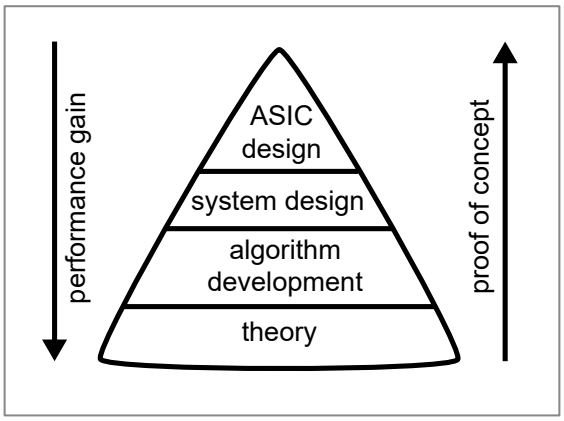


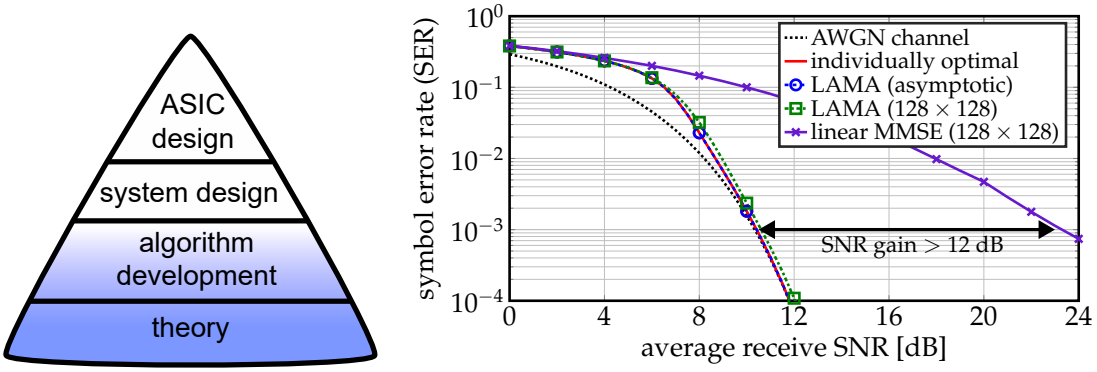

Figure 1.1: This thesis takes a holistic approach spanning theory, algorithm development, system design, and digital VLSI design to propose a solution for optimal data detection in practical massive MU-MIMO systems.

## 1.2 Contributions in Massive MU-MIMO

This thesis jointly investigates theory, algorithm development, system design, and application-specific integrated circuit (ASIC) design to enable optimal data detection in practical massive MU-MIMO systems. Fig. 1.1 illustrates the research approach pursued throughout this work: we develop novel algorithms and analyze their theoretical performance, present algorithm-level solutions to address implementation challenges, and design very-large scale integration (VLSI) architectures, and demonstrate the efficacy of the developed algorithms via real-world ASIC prototypes. Each of the following sections will detail the key contributions this thesis and list the associated publications.

### 1.2.1 Individually-Optimal Data Detection

Data detection in the massive MU-MIMO uplink (UEs transmit data to the BS) is among the most critical tasks from a spectral efficiency, hardware com-




(a) Contributions of Chapter 3

(b) Performance of LAMA in 128 × 128 i.i.d. Rayleigh massive MU-MIMO system and QPSK constellation

Figure 1.2: Contributions of Chapter 3: We develop a computationally efficient algorithm, referred to as LAMA which is able to achieve the error-rate performance of the so-called individually-optimal data-detector under certain assumptions on the MIMO channel matrix and the constellation alphabet.

plexity, and energy-efficiency perspective. However, solving the optimal data detection problem in MU-MIMO wireless systems with a large number of UEs entails prohibitive computational complexity. Although computationally efficient optimal data detection methods exist for small-scale MIMO systems, such as sphere-decoding (SD) based methods [2–4], their average complexity stills scales exponentially in the number of UEs. Consequently, such methods cannot be used. In order to reduce complexity of optimal data detection, a variety of sub-optimal detection algorithms have been proposed in the literature [5–8].

In this chapter of the thesis, we develop a novel, computationally efficient algorithm, referred to as LAMA, which is short for large MIMO approximate message passing). LAMA is able to achieve the error-rate performance of the individually-optimal (IO) data-detector under certain assumptions on the MIMO channel matrix and the constellation alphabet. The proposed method builds upon complex-valued AMP [9] and (real-valued) Bayesian AMP [1], initially proposed for sparse signal recovery. This thesis first develops the complex

Bayesian AMP (cB-AMP) framework along with its complex state evolution (cSE) framework. Then, cB-AMP is specialized to data detection in massive MU-MIMO systems, resulting in the LAMA algorithm. Fig. 1.2 shows that LAMA approaches the IO symbol error-rate (SER) performance by less than 0.2 dB SNR, whereas the reference linear minimum-mean squared error (MMSE) detector suffers a substantial SNR performance loss for a 128 BS antenna, 128 UE massive MU-MIMO system. The key contributions are summarized as follows:

- We propose the LAMA algorithm for data detection in massive MU-MIMO systems.
- We study the theoretical properties of LAMA in the massive MU-MIMO limit, i.e., when $\beta \to 0$ and the channel matrix is distributed i.i.d. Gaussian, and show that for such a scenario, simple low-complexity algorithms achieve IO performance.
- We demonstrate that the SE recursions of LAMA are identical to the fixed-point equations that predict the optimal multiuser efficiency developed in [10–12].
- We develop conditions for which LAMA achieves the same error-rate performance as the IO data detector.
- We derive exact recovery thresholds (ERTs), for which LAMA perfectly recovers signals from PAM, PSK, and QAM alphabets in noiseless scenarios.
- We investigate the achievable rates and error-rate performance of LAMA for PAM, PSK, and QAM constellations, and analyze the impact of the UE-to-BS ratio $\beta = U/B$.
- We characterize the performance/complexity trade-off of LAMA and show

that only a few algorithm iterations are sufficient to achieve near-IO performance.

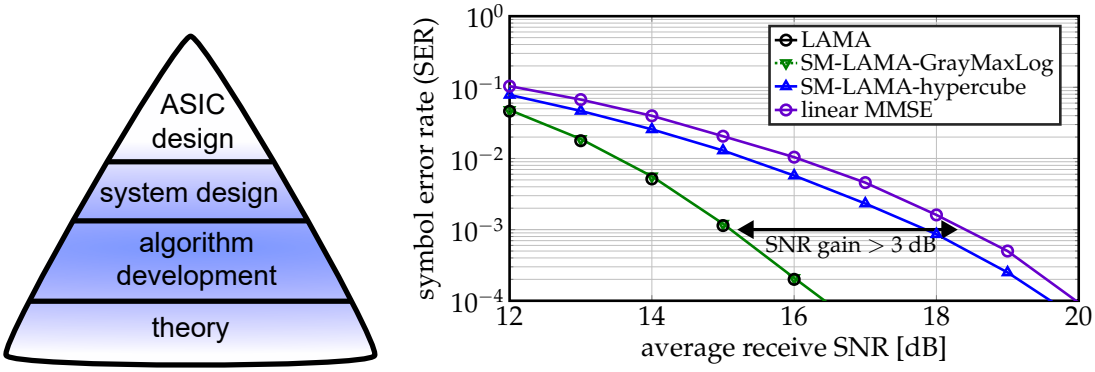
- We discuss the efficacy and limits of the proposed LAMA algorithms in practical (finite-dimensional) large-MIMO systems and provide corresponding numerical results.

The publications describing parts of the results in this chapter:

1. **C. Jeon**, R. Ghods, A. Maleki, and C. Studer, "Optimal data detection in large MIMO," *submitted to IEEE Transactions on Information Theory*, 2018
2. **C. Jeon**, R. Ghods, A. Maleki, and C. Studer, "Optimality of large MIMO detection via approximate message passing," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 1227-1231
3. R. Ghods, **C. Jeon**, A. Maleki, and C. Studer, "Optimal large-MIMO data detection with transmit impairments," *Proc. Allerton Conference on Communication, Control, and Computing*, Sept. 2015, pp. 1211-1218

### 1.2.2 Mismatched Data Detection

For practical, finite-dimensional massive MU-MIMO systems, LAMA enables near-IO performance at low computational complexity. Despite these advantages, LAMA requires repeated computations of transcendental functions that exhibit a high numerical precision. These computations render the design of corresponding high-throughput hardware designs that deploy finite precision (fixed-point) arithmetic a challenging task. Therefore, one often resorts to nu-




(a) Contributions of Chapter 4

(b) Performance of M-LAMA in $128 \times 64$ i.i.d. Rayleigh massive MU-MIMO system and 16-QAM constellation

Figure 1.3: Contributions of Chapter 4: We show that by carefully selecting a mismatched signal prior, we can not only prevent the need for transcendental functions and reduce the dynamic range, but also reduce the computational complexity.

merical approximations, which reduce the precision requirements and improve hardware efficiency.

In this chapter of the thesis, we present a theoretical framework to analyze and benchmark such approximate solutions that yield hardware friendly algorithms. To do so, we use the framework of mismatched data detection, when the prior used for detection differs from the true prior distribution. We show that by carefully selecting a mismatched signal prior, one can not only prevent the need for transcendental functions and reduce the dynamic range, but also reduce the computational complexity of data detection. We develop a mismatched version of the complex Bayesian approximate message passing (cB-AMP) presented in the thesis that includes a tuning stage to minimize the performance loss caused by a mismatch in the signal prior. To enable a precise performance analysis in the large-system limit, we propose the associated mismatched state-evolution (SE) framework. We then propose a mismatched version of the LAMA algorithm (short M-LAMA) and apply our framework to mismatched data detection in

large MIMO systems for two cases: (i) general case and (ii) PAM/QAM constellations. Fig. 1.3 shows that suboptimal mismatched LAMA (SM-LAMA) achieves identical SER performance as LAMA for a 128 BS antenna, 64 UE massive MU-MIMO system for 16-QAM constellation in a hardware-friendly way. The key contributions are summarized as follows:

- We present mismatched LAMA (M-LAMA) with a Gaussian prior for the general case in massive MU-MIMO. We present the corresponding SE analysis for M-LAMA which coincides exactly to the classical results from linear data detectors [13].
- We present M-LAMA with a uniform prior for PAM/QAM constellation. We show that the asymptotic performance of M-LAMA coincides exactly to that given by convex methods in [14, 15]. In addition, our analysis is constructive in a sense that M-LAMA achieves identical performance to that given by convex-based methods in a computationally efficient manner.
- We also present M-LAMA with a Gray-coding based approximation for PAM/QAM constellation. We show that M-LAMA achieves minimal loss in performance compared to the optimal LAMA algorithm, but in a hardware-friendly and low-complexity way.
- We present optimal and suboptimal tuning strategies and the associated computational complexity trade-offs for all the proposed mismatched prior distributions.
- We analyze the performance of the resulting algorithms in the large-system limit and support our findings via numerical simulations in finite-dimensional systems. These results demonstrate that carefully-selected

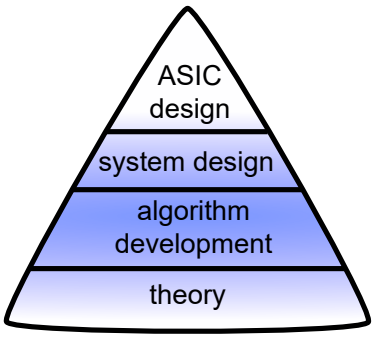
mismatched priors enable near-IO performance in finite-dimensional systems and in a hardware-friendly way.

The publications describing parts of the results in this chapter:

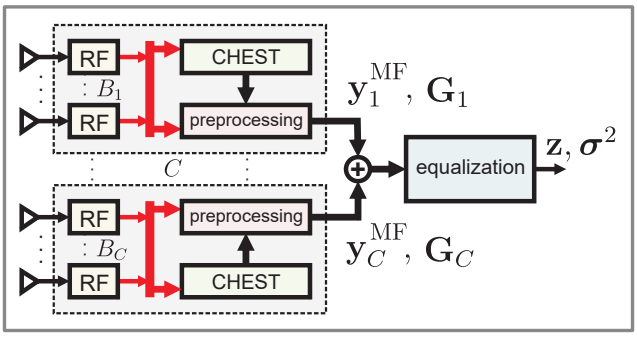
1. **C. Jeon**, A. Maleki, and C. Studer, "Mismatched data detection in large MIMO systems," *in preparation for IEEE Transactions on Signal Processing*, 2018
2. **C. Jeon**, A. Maleki, and C. Studer, "On the performance of mismatched data detection in large MIMO systems," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jul. 2016, pp. 180-184

### 1.2.3 Decentralized Data Detection

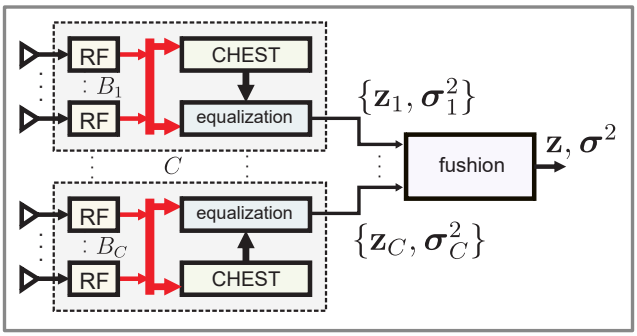
Although massive MU-MIMO promises orders-of-magnitude improvements in terms of spectral and energy efficiency, the large number of antennas at the BS causes significant challenges when implementing this technology. One of the most prominent challenges is the excessively high amount of fronthaul data that must be transferred from the radio-frequency (RF) antenna units at the BS antenna array to the baseband processing unit (BBU) [16–19]. Such high data rates not only exceed the bandwidth of existing high-speed interconnect standards, such as the common public radio interface (CPRI) [20], but also approach the limits of existing chip input/output (I/O) interfaces in terms of bandwidth and power dissipation [21]. Furthermore, traditional data-detection algorithms that achieve near-optimal spectral efficiency in the MU-MIMO uplink [6], such as zero-forcing (ZF) and linear minimum mean-square error (L-MMSE)-based




(a) Contributions of Chapter 5




(b) Partially decentralized equalization architecture




(c) Fully decentralized equalization architecture

Figure 1.4: Contributions of Chapter 5: We propose two distinct feedforward architectures for partially decentralized (PD) and fully decentralized (FD) equalization, which mitigate the interconnect, I/O, latency, and computation bottlenecks in massive MU-MIMO.

equalization, rely on centralized processing in a single computing fabric, which results in excessively high complexity for large antenna arrays [17, 22].

In this chapter of the thesis, we propose two distinct feedforward architectures for partially decentralized (PD) and fully decentralized (FD) equalization, which mitigate the interconnect, I/O, latency, and computation bottlenecks, shown in Fig. 1.4. For both of these architectures, we investigate the efficacy of matched filter (MF), ZF, L-MMSE, and LAMA for equalization for the massive MU-MIMO uplink. We present asymptotic analysis that discuss the performance and interconnect trade-offs between the proposed two architectures. Our results demonstrate that feedforward equalization enables scalable decentralized solutions for massive MU-MIMO systems with hundreds of antenna elements,

and incurs no or only a small loss in post-equalization SINR and error-rate performance compared to that of centralized solutions. The key contributions are summarized as follows:

- We develop a framework that enables a precise analysis of the post-equalization signal-to-noise-and-interference-ratio (SINR) of decentralized equalization with feedforward architectures in the large-system limit.
- We show that the partially-decentralized feedforward architecture achieves the same SINR performance as centralized solutions for equalization with MF, ZF, L-MMSE, and LAMA-PD.
- We show that the FD feedforward architecture is able to provide near-optimal SINR performance, but further reduces the interconnect and I/O bandwidths.
- We analyze optimal antenna partitioning strategies that maximize the SINR for the FD architecture.
- We conduct error-rate simulations for a realistic 3GPP long-term evolution (LTE)-like massive MU-MIMO system that support our theoretical findings.

The publications describing parts of the results in this chapter:

1. **C. Jeon**, K. Li, J. R. Cavallaro, and C. Studer, "Decentralized equalization with feedforward architectures for massive MU-MIMO," *submitted to IEEE Transactions on Signal Processing*, 2018
2. **C. Jeon**, K. Li, J. R. Cavallaro, and C. Studer, "On the achievable rates of decentralized equalization in massive MU-MIMO systems," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 1102-1106

3. K. Li, **C. Jeon**, J. R. Cavallaro, and C. Studer, "Decentralized equalization for massive MU-MIMO on FPGA," *Proc. Asilomar Conference on Signals, Systems, and Computers*, Oct. 2017, pp. 1532-1536
4. K. Li, **C. Jeon**, J. R. Cavallaro, and C. Studer, "Feedforward architectures for decentralized precoding in massive MU-MIMO systems," *Proc. Asilomar Conference on Signals, Systems, and Computers*, 2018 (**2nd Place at the Student Paper Contest)**

### 1.2.4 Efficient Preprocessing

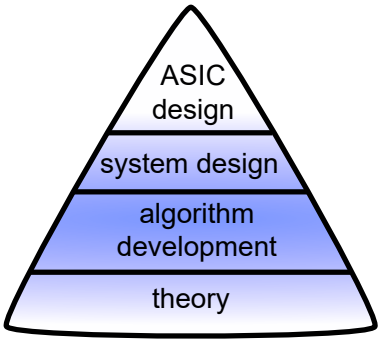
The computational complexity of most existing linear and non-linear data-detection algorithms in massive MU-MIMO systems is often dominated by the computation of the Gram matrix [22, 23]. The complexity of Gram matrix computation is orders-of-magnitude higher in wideband systems that use orthogonal frequency-division multiplexing (OFDM) or single-carrier frequency-division multiple access (SC-FDMA), in which this matrix must be computed for each active subcarrier (i.e., subcarriers used for pilots and data transmission) [22].

In this chapter of the thesis, inspired by exact, interpolation-based matrix computation algorithms in [24–27] for small-scale wideband MIMO systems, we propose novel algorithms for *approximate* Gram matrix computation in massive MU-MIMO systems. Instead of performing a brute-force (BF) computation of the Gram matrix for all subcarriers or using exact interpolation schemes, we propose two simple, yet efficient approximate interpolation methods that perform close to that of an exact BF computation at only a fraction of the computational complexity. In addition, the proposed approximate interpolation methods are more robust

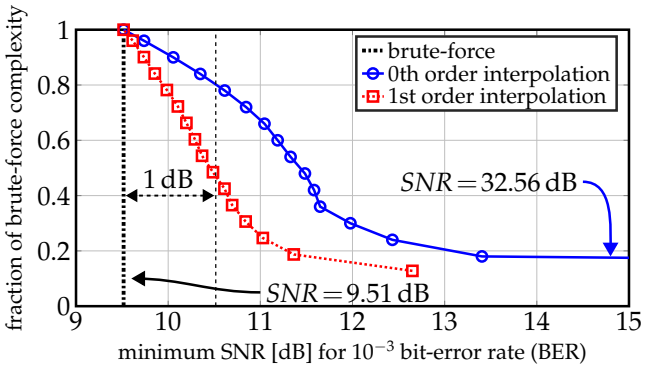
to channel-estimation errors and receive-side antenna correlation than exact interpolation methods. Fig. 1.5 shows the trade-off between minimum SNR required to achieve $10^{-3}$ bit-error rate (BER) in a $128 \times 8$ i.i.d. Rayleigh massive MU-MIMO system and 16-QAM constellation with channel estimation errors. We see that the 1st order interpolation method approaches the SNR performance of the brute-force method by 1dB with only 45% of the complexity. The key contributions are summarized as follows:

- We study exact methods and establish the minimum number of Gram matrix base-points that are required for exact interpolation.
- We show that channel-hardening in massive MU-MIMO enables approximate interpolation schemes that achieve near-exact error-rate performance, even with strong undersampling in the frequency domain.
- We provide analytical results that characterize the approximation errors of the proposed interpolation methods depending on the channel's delay spread and the antenna configuration.
- We also derive exact mean-squared error (MSE) expressions of our approximate interpolation algorithms for imperfect channel-state information (CSI) and BS-antenna correlation.
- We characterize the trade-offs between computational complexity and error-rate performance in realistic massive MU-MIMO-OFDM systems, and we demonstrate the robustness of our approximate interpolation methods for realistic scenarios with imperfect CSI and BS-antenna correlation.

The publication describing parts of the results in this chapter:




(a) Contributions of Chapter 6



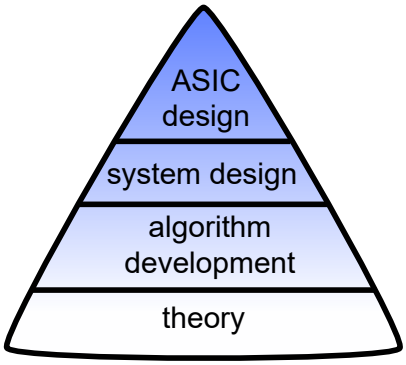

(b) 1st order interpolation method approaches the SNR performance of the brute-force method by 1 dB with only 45% of the complexity

Figure 1.5: Contributions of Chapter 6: We propose two simple, yet efficient approximate interpolation methods for computing the Gram matrix that perform close to that of an exact computation at only a fraction of the computational complexity.

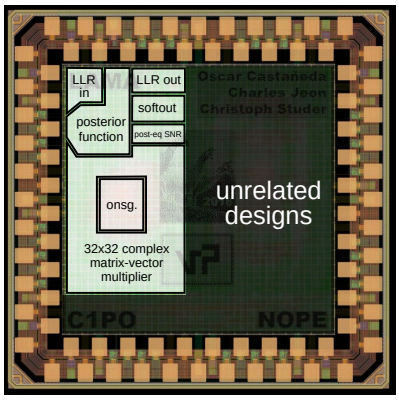
1. **C. Jeon**, Z. Li, and C. Studer, "Approximate Gram-matrix interpolation for wideband massive MU-MIMO systems," *submitted to IEEE Transactions on Wireless Communications*, 2018

### 1.2.5 ASIC Design of Optimal Data Detection

To enable high-throughput data detection performance for massive MU-MIMO systems, a variety of low-complexity algorithms (see [7, 28–30] and references therein) as well as the corresponding hardware designs in field-programmable gate array (FPGA) [8, 22, 31–34] and application-specific integrated circuit (ASIC) [35–37] designs have been proposed in the existing literature. Majority of these algorithms, and their hardware designs rely on (approximate) linear data detection, which are known to scale poorly in massive MU-MIMO uplink when the number of transmit antennas is similar to that of receive antennas at the BS [22]. Moreover, these algorithms are designed to operate on channels that have i.i.d.




(a) Contributions of Chapter 7




(b) ASIC Design of a LAMA

Figure 1.6: Contributions of Chapter 7: We provide the first, to the best of our knowledge, 32-user soft-input and soft-output near-optimal data detector for massive MU-MIMO. We prototype the LAMA algorithm and show that near-optimal data detection is possible in practical massive MU-MIMO systems by a joint optimization of theory, algorithm, system, and ASIC design.

Rayleigh fading, which in general, does not hold in practical massive MU-MIMO systems.

In this chapter of the thesis, we propose an ASIC implementation of a massive MU-MIMO data detector that is capable of soft-input soft-output (SISO) capabilities for iterative MIMO data detection and decoding. Our implementation is based on the LAMA (short for large MIMO approximate message passing (AMP)) algorithm [38, 39] that was shown achieve optimal error-rate performance in the asymptotic regime, and under i.i.d. Rayleigh fading. Since these two assumptions are not satisfied in general, we extend the LAMA algorithm, and present numerous algorithmic- and hardware-level optimizations to arrive at a high-throughput and low-area ASIC design, while maintaining a robust near-optimal error-rate performance in practical massive MU-MIMO systems. Our key contribution is as follows:

- We provide an ASIC implementation of the LAMA algorithm [38, 39] that

can provably achieve optimal error-rate performance in the asymptotic regime.

- We present algorithmic-, architecture-level, and hardware-level optimizations that leads to an efficient ASIC implementation.
- We compare our LAMA results to existing data-detection algorithms and showcase the significant gains in spectral efficiency under realistic channel models.

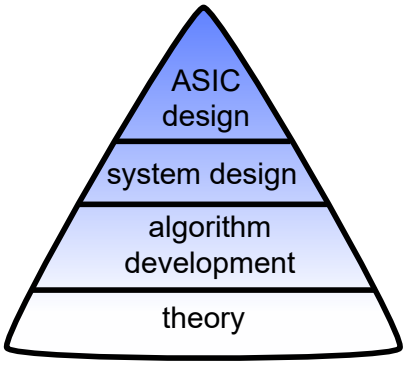
The publication describing parts of the results in this chapter:

1. **C. Jeon**, O. Castañeda, and C. Studer, "ASIC implementation of a 32-user soft-input soft-output massive MU-MIMO detector," *in preparation for submission*, 2018

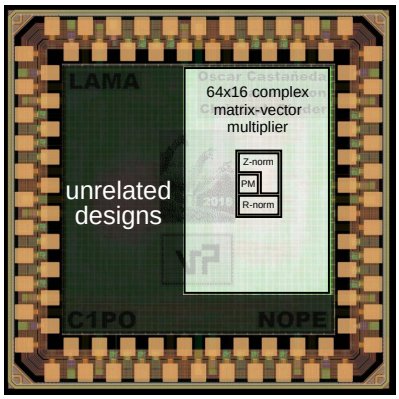
### 1.2.6 ASIC Design of a Nonparametric Equalizer

Low-complexity, but near-optimal data detection algorithms are a key to achieve high spectral efficiency in the fifth-generation wireless systems. However, practical massive MU-MIMO systems suffer—as do traditional MIMO systems—from real-world hardware impairments and model mismatches, e.g., amplifier nonlinearities, phase noise, quantization artifacts, and channel-estimation errors. Such system nonidealities are known to substantially reduce the performance of optimal data-detection algorithms unless one explicitly models these impairments and estimates the associated parameters [40].

In this chapter of the thesis, we develop a hardware architecture of the NOPE (short for nonparametric equalizer) algorithm proposed in [41] that requires low




(a) Contributions of Chapter 8




(b) ASIC Design of a NOPE

Figure 1.7: Contributions of Chapter 8: We provide the first VLSI design of NOPE algorithm [41]. Our results demonstrate that massive MU-MIMO has the unique potential to design parameter-free algorithms, such as NOPE, that perform on par with solutions that require accurate knowledge of critical system and model parameters.

complexity and is robust to system impairments and model mismatches. We generalize NOPE to practical channels and provide, to the best of our knowledge, its first VLSI design. Our results demonstrate that massive MU-MIMO has the unique potential to design parameter-free algorithms, such as NOPE, that perform on par with solutions that require accurate knowledge of critical system and model parameters. The key contributions of this chapter are summarized as follows:

- We propose a set of algorithm-level modifications that enable NOPE to operate on more realistic MU channels.
- We develop a VLSI architecture that relies on Cannon's algorithm [42] to achieve high throughput at low area.
- We show reference VLSI synthesis results in 28 nm CMOS for a 64 BS antenna, 16 UE massive MU-MIMO system.
- We compare NOPE to existing massive MU-MIMO equalizers requiring

knowledge of the signal and noise powers.

The publication describing parts of the results in this chapter:

1. **C. Jeon**, G. Mirza, R. Ghods, A. Maleki, and C. Studer, "VLSI design of a nonparametric equalizer for massive MU-MIMO," *Proc. Asilomar Conference on Signals, Systems, and Computers*, Oct. 2017, pp. 1504-1508
2. R. Ghods, **C. Jeon**, G. Mirza, A. Maleki, and C. Studer, "Optimally-tuned nonparametric linear equalization for massive MU-MIMO systems," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 2118-2122

## 1.3 Notation

Lowercase and uppercase boldface letters represent column vectors and matrices, respectively. For a matrix $\mathbf{H}$, we define its transpose and Hermitian to be $\mathbf{H}^{\mathrm{T}}$ and $\mathbf{H}^{\mathrm{H}}$, respectively. The $b$th row vector and $u$th column of the matrix $\mathbf{H}$ are denoted by $\mathbf{h}_b^{\mathrm{row}}$ and $\mathbf{h}_u^{\mathrm{col}}$ respectively, the entry on the $b$th row and $u$th column is $H_{b,u}$, and the $u$th entry of a vector $\mathbf{x}$ is $x_u$. For a $N$-dimensional vector $\mathbf{x}$, we define its complex conjugate by $\mathbf{x}^*$ and its $u$th entry by $x_u$. The $U \times U$ identity matrix is denoted by $\mathbf{I}_U$ and the $B \times U$ all-zeros matrix by $\mathbf{0}_{B \times U}$. The real and imaginary parts of scalars, vectors, and matrices are denoted by $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$, respectively. We use $\langle \cdot \rangle$ to represent the averaging operator $\langle \mathbf{x} \rangle = \frac{1}{N} \sum_{k=1}^{N} x_k$. Multivariate real-valued and complex-valued Gaussian probability density (pdf) functions are denoted by $\mathcal{N}(\mathbf{m}, \mathbf{K})$ and $\mathcal{CN}(\mathbf{m}, \mathbf{K})$, respectively, where $\mathbf{m}$ is

the mean vector and $\mathbf{K}$ the covariance matrix; $\mathbb{E}_X[\,\cdot\,]$ denotes expectation and $\mathbb{V}\mathrm{ar}_X[\,\cdot\,]$ denotes variance with respect to the pdf of the random variable $X$.

## 1.4 Full Publication List

- **C. Jeon**, R. Ghods, A. Maleki, and C. Studer, "Optimal data detection in large MIMO," *submitted to IEEE Transactions on Information Theory*, 2018
- **C. Jeon**, K. Li, J. R. Cavallaro, and C. Studer, "Decentralized equalization with feedforward architectures for massive MU-MIMO," *submitted to IEEE Transactions on Signal Processing*, 2018
- **C. Jeon**, Z. Li, and C. Studer, "Approximate Gram-matrix interpolation for wideband massive MU-MIMO systems," *submitted to IEEE Transactions on Wireless Communications*, 2018
- **C. Jeon**, O. Castañeda, and C. Studer, "ASIC implementation of a 32-user soft-input soft-output massive MU-MIMO detector," *in preparation for IEEE Journal of Solid-State Circuits*, 2018
- **C. Jeon**, A. Maleki, and C. Studer, "Mismatched data detection in large MIMO systems," *in preparation for IEEE Transactions on Signal Processing*, 2018
- R. Ghods, **C. Jeon**, A. Maleki, and C. Studer, "Optimal data detection and signal estimation with input noise," *in preparation for IEEE Transactions on Signal Processing*, 2018
- S. Jacobsson, O. Castañeda, **C. Jeon**, G. Durisi, and C. Studer, "Nonlinear precoding for phase-quantized constant-envelope massive MU-MIMO-

OFDM," *Proc. IEEE International Conference on Telecommunications (ICT)*, Jun. 2018, pp. 367-372

- K. Li, **C. Jeon**, J. R. Cavallaro, and C. Studer, "Feedforward architectures for decentralized precoding in massive MU-MIMO systems," *Proc. Asilomar Conference on Signals, Systems, and Computers*, 2018 (**2nd Place in the Student Paper Contest**)
- **C. Jeon**, C. Studer, M. Wu, and C. Dick, "System and method for downlink processing in communication systems," *United States Patent and Trademark Office (USPTO)*, 9876657, 2017
- **C. Jeon**, G. Mirza, R. Ghods, A. Maleki, and C. Studer, "VLSI design of a nonparametric equalizer for massive MU-MIMO," *Proc. Asilomar Conference on Signals, Systems, and Computers*, Oct. 2017, pp. 1504-1508
- K. Li, **C. Jeon**, J. R. Cavallaro, and C. Studer, "Decentralized equalization for massive MU-MIMO on FPGA," *Proc. Asilomar Conference on Signals, Systems, and Computers*, Oct. 2017, pp. 1532-1536
- R. Ghods, **C. Jeon**, G. Mirza, A. Maleki, and C. Studer, "Optimally-tuned nonparametric linear equalization for massive MU-MIMO systems," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 2118-2122
- **C. Jeon**, K. Li, J. R. Cavallaro, and C. Studer, "On the achievable rates of decentralized equalization in massive MU-MIMO systems," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 1102-1106
- **C. Jeon**, A. Maleki, and C. Studer, "On the performance of mismatched data detection in large MIMO systems," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jul. 2016, pp. 180-184

- R. Ghods, **C. Jeon**, A. Maleki, and C. Studer, "Optimal large-MIMO data detection with transmit impairments," *Proc. Allerton Conference on Communication, Control, and Computing*, Sept. 2015, pp. 1211-1218
- **C. Jeon**, R. Ghods, A. Maleki, and C. Studer, "Optimality of large MIMO detection via approximate message passing," *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 1227-1231

CHAPTER 2

# MASSIVE MU-MIMO: BASICS AND STATE-OF-THE-ART

We will first motivate the need for massive MU-MIMO by discussing the exponential growth of data consumption in mobile traffic. We will then discuss the existing state-of-the-art MIMO, as well as massive MU-MIMO technology. In addition, we detail the system model and the wireless system that we will assume throughout this thesis.

## 2.1 Motivation for Massive MU-MIMO: Explosive Growth of Data Consumption

Fig. 2.1 shows the global monthly cellular traffic trend, which shows the exponential growth of data which dwarfs traffic incurred by voice, which stays constant [43]. This paradigm shift of voice-based mobile communication to data-based mobile communication is also fueled by an exponential increase number of wireless-capable devices, ranging from portable computers to Internet-of-Things (IoT) [44]. The driving force behind data consumption is more users streaming high-resolution video content in their mobile platform, and the trend of moving to 720p or 1080p resolution from standard video resolution of 480p [43]. The exponential growth of cellular data from video traffic is forecasted to continue to grow at 45 percent annually through 2023 [43] to account for 73 percent of all mobile data traffic. Moreover, new trends of new video technology is expected to significantly impact data traffic consumption. One example is YouTube 360-degree video, which consumes more than $4\times$ bandwidth compared to standard video at the same resolution [43].

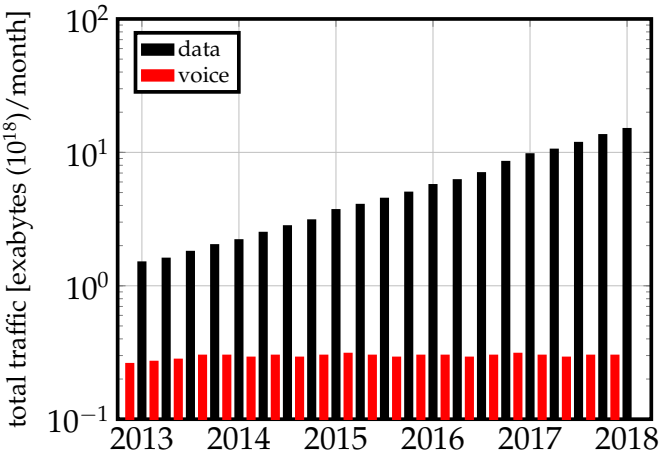



Figure 2.1: Measured global monthly cellular traffic data/voice [43]: Data growth is increasing exponentially whereas voice stays relatively constant.

Unfortunately, however, state-of-the-art systems are already at its limits in terms of throughput and therefore, unable to cope with the increase in data consumption [43]. As a result, to provide a good quality-of-service for all UE while simultaneously connecting all devices, new wireless communication standards are required to provide higher spectral efficiency. One viable solution to improve spectral efficiency for multiple UEs is massive MU-MIMO, which we will describe next.

## 2.2 MIMO and Massive MU-MIMO Wireless Technology

Multiple-input multiple-output (MIMO) is a physical layer technique that provides an increase in spectral efficiency without additional increase in transmit power and/or bandwidth. First proposed for consumer technology in IEEE 802.11n-2009 [47], MIMO technology has been incorporated to virtually all currently utilized wireless systems, which include LTE and LTE advanced [48–50], WiMAX, and IEEE 802.11ac and 802.11ax [51]. The key idea is to utilize multiple

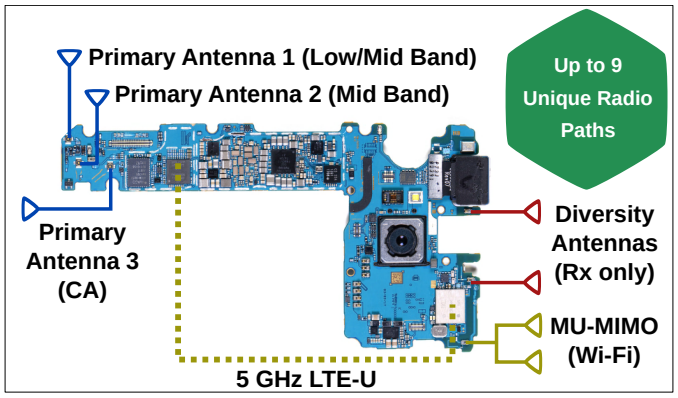



Figure 2.2: Samsung Galaxy S8 utilizes multiple antennas to achieve downlink speeds of 1 Gigabit per second [45, 46].

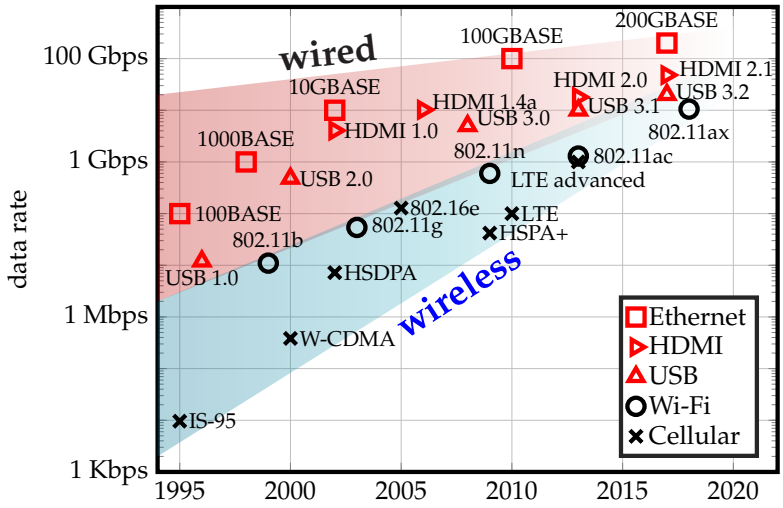



Figure 2.3: Trends and data rates of modern wired and wireless technology: Slope of the date rate of wireless is higher than that given by wired technology.

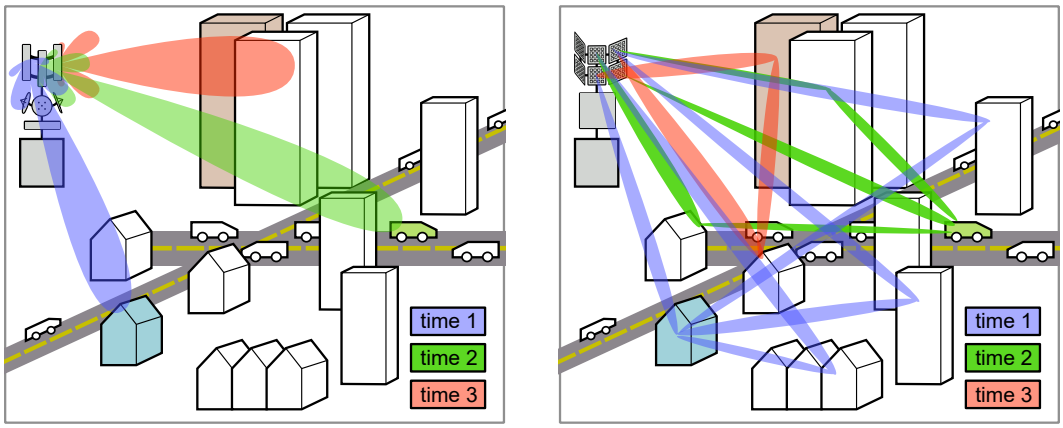
antennas at the transmitter and/or receiver simultaneously to achieve higher throughput and/or improving reliability. Since these additional antennas allow multiple data streams to be utilized independently, in general, MIMO offers significantly increased throughput compared to single-input single-output system. Fig. 2.2 shows an example of a Samsung Galaxy S8 smartphone, which utilizes multiple antenna simultaneously to achieve maximum speed of 1 Gigabit per second. In addition, the smartphone has two diversity antennas which help to improve reliability.

Fig. 2.3 shows data rates for commonly used wired and wireless systems for consumer technology. We observe that the gap between wired and wireless communications is diminishing, and it is predicted that wireless will eventually provide higher throughput than wired technology [52]. The trend is also reflected in Edholm's law [52], which states that telecommunications data rates are doubling every 18 months, which is exactly identical to the transistor scaling in Moore's law.

### 2.2.1 Massive MU-MIMO

To address the exponential increase in data without simply increasing the power nor allocating more bandwidth, Marzetta proposed a "massive" or large-scale MIMO system where the number of antennas are scaled up even at a greater magnitude [53]. An analysis of scaling the BS antennas to infinity, while keeping the number of user antennas constant, showed that the effect of fading disappears completely. This effect was also known as "wires-in-the-air" [54]. Moreover, in the asymptotic regime simple matched-filter equalization was proved to be optimal (see Section 3.3.4 for a detailed discussion).

The largest difference between massive MIMO and current MIMO is the number of antennas at the BS. Current state-of-the art MIMO standards such as LTE-Advanced [48] and/or IEEE 802.11ax [51] supports up to 8 BS and 8 user equipment (UE), whereas massive MIMO envisions the usage of hundreds or even thousands of antennas while only increasing the UE to the order of tens. Due to the order-of-magnitude increase in the antennas at the BS, massive MIMO is also commonly referred as large-scale MIMO or very-large MIMO [5]. We




(a) Conventional MIMO (b) Massive MIMO

Figure 2.4: Illustration of conventional MIMO and massive MIMO: Unlike conventional MIMO, massive MIMO allows fine-grained beamforming that simplifies the UE equalization and data detection procedure.

highlight some of the advantages of massive MIMO:

1. **Uplink–Fading is significantly reduced in Massive MIMO:** In the uplink (UL), where UE is communicating to the BS, the BS performs equalization and detection on the signals sent by the UE. Although the wireless fading is random due to the propagation environment, combining the received signals coherently significantly reduce fading, and virtually eliminate fading in the asymptotic limit [6]. This phenomena is also known as "channel hardening" [5, 55, 56]. Since fading is reduced significantly, equalization processing can be made simpler at the BS.

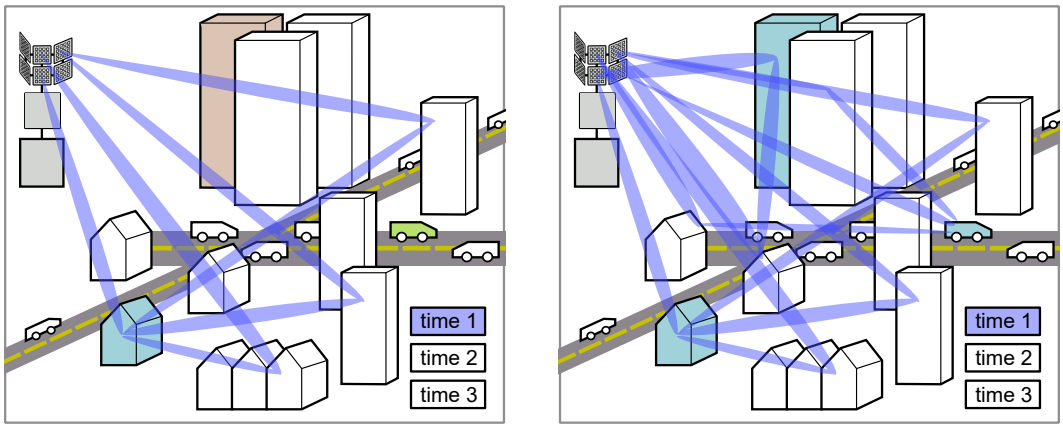
2. **Downlink–Fine-grained beamforming:** In the downlink (DL), where BS is communicating to the UEs, the BS can utilize the additional antennas to perform preprocessing the UE's data signal. The additional antennas can be used to constructively add the signal directed to each UE increasing the received signal strength, while destructively removing the inter-UE interference. This procedure, known as precoding, can be done in the digital

domain without additional analog circuitry. Precoding couples well with massive MIMO as larger number of antennas provide a greater spatial resolution. As a result, precoding provides the UE with its corresponding data at a significantly lower interference compared to existing MIMO systems. Fig. 2.4 shows a the prime difference between conventional MIMO and massive MIMO. The BS can utilize the wireless channel knowledge to perform much finer beamforming Fig. 2.4(b) compared to that of conventional MIMO Fig. 2.4(a).

3. **Reduced power consumption at UE:** In the uplink, fading is reduced by fine-grained beamforming at the BS, utilizing the large number of arrays to perform coherent combining. Therefore, the UE can operate at a lower signal-to-noise ratio and hence, reducing transmit signal power. In the downlink, precoding pushes the computational complexity to the BS so that UE only has to perform simple equalization and data detection. This shift of computational burden decreases the hardware cost at the UE due to simpler communication hardware and reduced power consumption.

On top of providing higher link reliability and reduced power consumption at the UE, massive MIMO can provide significantly higher spectral efficiency compared to existing MIMO systems by serving multiple users simultaneously at the same time-frequency resource.

Fig. 2.5 shows how multi-user MIMO can bring significantly improved spectral efficiency. In single-user MIMO shown in Fig. 2.5(a), a total of three users are communicating from the BS over three time slots, where we show only for time slot 1. While each UE can have multiple antennas, each UE occupies the whole time slot (blue, green, red) for communication. However, in MU-MIMO




(a) Single-user massive MIMO (b) Multi-user massive MIMO

Figure 2.5: Illustration of single-user and multi-user massive MIMO: Multi-user massive MIMO can bring significantly increased spectral efficiency over conventional state-of-the-art MIMO technology by serving multiple UEs in the same time-frequency resource.

shown in Fig. 2.5(b), multiple UEs (each with possibly multiple antennas) can be served *simultaneously* in a same time slot. In the uplink, the BS performs coherent equalization for the UE's transmit signals; in the downlink, the BS performs precoding for beamforming to remove inter-user interference. Since this procedure can be done on a per-time-slot basis, the MU-MIMO can serve multiple users over multiple time-slots. As shown in Fig. 2.5(b), the three UEs are served in a single time slot, which provides much greater spectral efficiency compared to one UE served for a single-user MIMO.

## 2.3 Introduction to Wireless Systems

In this section, we provide a brief overview of a wireless system and the associated challenges compared to wired communications. Unlike in wired communications, in which the medium of communication is done by a physical cable such as Ethernet or HDMI, the medium of communications in wireless is carried out

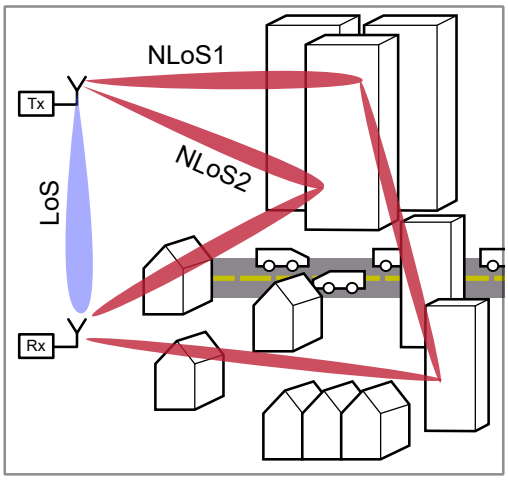
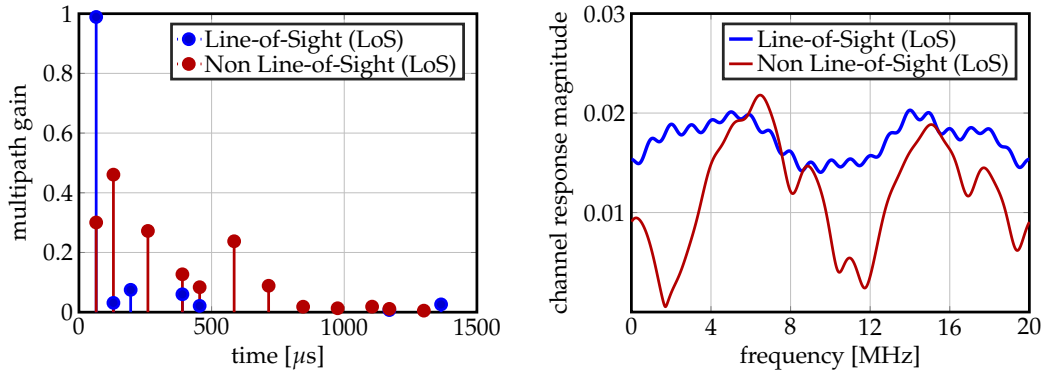



Figure 2.6: Illustration of fading in a wireless system: Wireless system may have both Line-of-Sight and Non-Line-of-Sight components




(a) time-domain impulse response (b) frequency-domain response

Figure 2.7: LoS and NLoS propagation effects and its impact on frequency-selective fading.

through electromagnetic waves generated by the radio-frequency (RF) chain. The transmitter generates the electromagnetic waves and propagates it through the air from its transmit antenna. The transmitted wave experiences different scattering and diffraction in paths between the transmitter and receiver, which results in a superposition of delayed and attenuated versions of the original waveform with different magnitude scaling arriving at the receiver antenna. Such effects are referred to as multi-path fading [57]. In general, we refer Line-of-Sight (LoS) path as the direct path, and Non-Line-of-Sight (NLoS) as all the diffracted paths

from the transmitter to receiver; we refer to all these multi-path effects as the channel between the transmitter and the receiver. An illustration of a wireless system is shown in Fig. 2.6. The transmitter (Tx) sends a wave to the receiver (Rx). The wave experiences three different paths comprising one LoS path and two NLoS paths.

The multi-path fading in the wireless channel in time domain brings effects in the frequency domain for both LoS and NLoS. Fig. 2.7(a) shows the impulse response for both LoS and NLoS scenario in a typical urban environment for Long-Term Evolution (LTE). We can observe that the LoS scenario has most of the energy concentrated in the first time step, whereas for NLoS scenario has energy scattered around multiple time steps. The differences in the delays for these two scenarios bring forth effects in the frequency domain, shown in Fig. 2.7(b). We observe that LoS has a slower varying magnitude compared to that given by NLoS. In addition of these fading scenarios, there are additional effects that affect the wireless channel, which include mobility of both transmitter and receiver. All of the aforementioned fading effects make wireless communications very challenging, and thus makes equalization and data detection an essential task to ensure reliable and robust communication.

## 2.4 System Model

We now describe the system model that we will use throughout the thesis. We consider an uplink MU-MIMO system where we have $U$ users each with a single antenna, and $B$ receive antennas at the base-station. A visual representation of the system is shown in Fig. 2.8. Contrary to the existing literature on classic

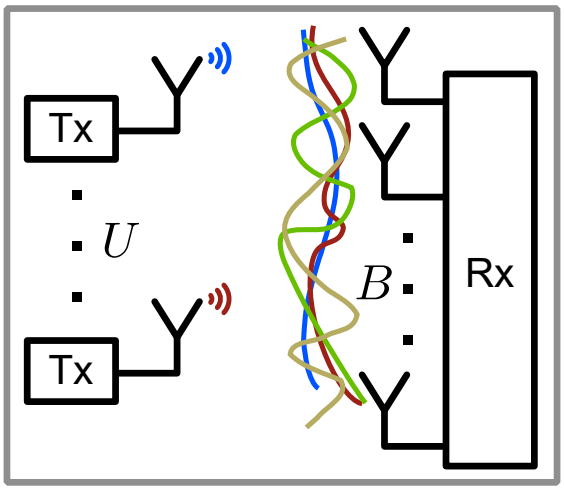



Figure 2.8: Input-output system model for the MU-MIMO uplink

MIMO, we do not place any restrictions on the relationship between the number of user and BS antennas; we will discuss regimes when we have more users than BS antennas, i.e., overloaded systems.

The input-output relation of the MIMO channel is modeled as:

$$\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}, \tag{2.1}$$

where $\mathbf{s}_0$ is the $U$-dimensional symbol sent from the transmitter, $\mathbf{y}$ is the $B$-dimensional vector received at the BS, and $\mathbf{H}$ is the complex-valued $B \times B$ channel matrix that represents the wireless fading characteristics, which includes the different transmit powers for each individual UE. We model the $B$-dimensional thermal noise $\mathbf{n}$ at the receiver i.i.d. zero mean circularly symmetric complex Gaussian distribution with variance $N_0$, i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{\times 1}, N_0\mathbf{I}_B)$. Throughout the thesis, we assume the BS has perfect knowledge of the channel[1] where as the UE does not have any knowledge of the channel. Since the transmitted symbol $\mathbf{s}_0$ comes from a discrete constellation set $\mathcal{O}^U$, e.g., phase shift keying (PSK) or pulse amplitude modulation (PAM), we model the prior distribution for each

[1] in practice, channel is obtained through pilot symbols

$u$th UE as

$$p(s_u) = \sum_{a_u \in \mathcal{O}} p_{a_u} \delta(s_u - a_u), \quad \sum_{a_u \in \mathcal{O}} p_{a_u} = 1,$$

where $\delta(\,\cdot,)$ is the Dirac delta function. For the case of uniform priors, we have $p(s_u) = \frac{1}{|\mathcal{O}|} \sum_{a \in \mathcal{O}} \delta(s_u - a)$. The signal-to-noise ratio (SNR) that we use throughout the thesis is the average SNR per receive antenna defined by

$$\mathsf{SNR} = \frac{\mathbb{E}_{\mathbf{H},\mathbf{s}} \left[ \|\mathbf{Hs}\|^2 \right]}{\mathbb{E}_{\mathbf{n}} \left[ \|\mathbf{n}\|^2 \right]}.$$

## 2.5 The Basics of MIMO Data Detection

We will start by reviewing data detection methods in MIMO. We will first formulate the optimal data detection problem and discuss why it becomes intractable in large or massive MIMO systems. We will then present heuristic linear data detection algorithms that are commonly used in MIMO, as well as massive MIMO.

### 2.5.1 Optimal Data Detection

For the system model presented in Fig. 2.8, we will first define two data-detection problems, which are jointly optimal (JO) and individually optimal (IO). We will formulate the optimization problem and discuss why solving them are a non-trivial task.

**Jointly Optimal Data Detection**

The JO problem, more commonly known as the maximum a-priori (MAP) problem [57,58], seeks to minimize the error-rate pertaining to the symbol of *all* users. Mathematically, the estimate $\hat{\mathbf{s}}^{\mathrm{JO}}$ is obtained by solving the problem:

$$\hat{\mathbf{s}}^{\mathrm{JO}} = \underset{\tilde{\mathbf{s}} \in \mathbb{C}^U}{\arg\max}\, p(\tilde{\mathbf{s}}|\mathbf{y}), \tag{2.2}$$

where $p(\tilde{\mathbf{s}}|\mathbf{y})$ is the posterior probability distribution for the system model in Fig. 2.8. By using Bayes rule, the problem in (2.2) can be factored into a MAP data detection problem, in which we will refer to as jointly optimal (JO) problem [58]:

$$\begin{aligned}
\hat{\mathbf{s}}^{\mathrm{JO}} &= \underset{\tilde{\mathbf{s}} \in \mathbb{C}^U}{\arg\max}\, p(\tilde{\mathbf{s}}|\mathbf{y}) \\
&= \underset{\tilde{\mathbf{s}} \in \mathbb{C}^U}{\arg\max}\, p(\mathbf{y}|\tilde{\mathbf{s}}) p(\tilde{\mathbf{s}}) \\
&\overset{(a)}{=} \underset{\tilde{\mathbf{s}} \in \mathbb{C}^U}{\arg\max} \left( -\frac{\|\mathbf{y} - \mathbf{H}\tilde{\mathbf{s}}\|_2^2}{N_0} + \log p(\tilde{\mathbf{s}}) \right),
\end{aligned}$$

where in $(a)$ we used the fact that $p(\mathbf{y}|\tilde{\mathbf{s}})$ follows complex Gaussian distribution with mean $\mathbf{H}\tilde{\mathbf{s}}$ and variance $N_0$ from Fig. 2.8, and that $\log(\,\cdot\,)$ is monotonically increasing. For the case of uniform likely transmit symbols, the (2.2) reduces to that given by maximum-likelihood problem [57,58], which is given as:

$$\hat{\mathbf{s}}^{\mathrm{JO}} = \underset{\tilde{\mathbf{s}} \in \mathcal{O}^U}{\arg\min}\, \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{s}}\|_2^2.$$

By construction of the JO problem (2.2), $\hat{\mathbf{s}}^{\mathrm{JO}}$ minimizes the vector error-rate, which is the probability that the symbol vector is different from the original transmitted symbol vector, i.e., $\Pr(\hat{\mathbf{s}}^{\mathrm{JO}} \neq \mathbf{s}_0)$.

The main difficulty of solving JO is currently no polynomial time algorithm is known; an exhaustive search over all the possible transmit symbol vectors entails exponential complexity. For a MIMO system with $U$ UEs, the complexity

of search operation is given by $|\mathcal{O}|^U$, which amounts to more than 16 million possible transmit symbol vectors for a system with 4 UEs and 64-QAM constellation. For a massive MU-MIMO system with tens of users, the computational complexity of solving the JO problem quickly becomes intractable.

**Individually Optimal Data Detection**

In practical multi-user wireless systems, it is important to maximize the performance of *each* UE individually. Thus, solving the JO problem may not be the optimal criterion for maximizing spectral efficiency in practical wireless systems. Therefore, the individually-optimal (IO) data detection problem estimates the most likely symbol for each individual UE separately, unlike the jointly-optimal problem, where we estimate the symbol most likely for all the users in a joint fashion. The IO problem is also commonly known as the element-wise MAP problem [58, 59], which is solved for each UE independently by solving the following optimization problem:

$$\hat{s}_u^{\text{IO}} = \underset{\tilde{s}_u \in \mathbb{C}}{\arg\max}\, p(\tilde{s}_u|\mathbf{y}), \quad u = 1, \ldots, U. \tag{2.3}$$

Similar to the simplification for the JO problem, we use Bayes rule to factor (2.3) into

$$\begin{aligned}
\hat{s}_u^{\text{IO}} &= \underset{\tilde{s}_u \in \mathbb{C}}{\arg\max}\, p(\tilde{s}_u|\mathbf{y}) \\
&= \underset{\tilde{s}_u \in \mathbb{C}}{\arg\max} \int_{\mathbb{C}^{U-1}} p(\mathbf{y}|\tilde{\mathbf{s}}) p(\tilde{\mathbf{s}}) \mathrm{d}\tilde{\mathbf{s}}^u \\
&= \underset{\tilde{s}_u \in \mathbb{C}}{\arg\max} \int_{\mathbb{C}^{U-1}} \exp\left( -\frac{\|\mathbf{y} - \mathbf{H}\tilde{\mathbf{s}}\|^2}{N_0} + \log p(\tilde{\mathbf{s}}) \right) \mathrm{d}\tilde{\mathbf{s}}^u,
\end{aligned} \tag{2.4}$$

where we introduced a shorthand notation $\mathrm{d}\tilde{\mathbf{s}}^u$ that so that the integral is computed over $\mathbb{C}^{U-1}$, leaving only the probability distribution of $u$th UE. In short,

solving the IO problem for the $u$th UE amounts to first marginalizing the joint probability distributions $p(\mathbf{y}|\tilde{\mathbf{s}})$ without the $u$ UE, and then picking the closest symbol $a_u \in \mathcal{O}$ that maximize this marginalized a-posteriori distribution. By construction of the IO problem, $\hat{\mathbf{s}}^{\text{IO}}$ minimizes the per-user symbol error-rate [58], which is the probability that each symbol is different than each UE's transmitted symbol, i.e., $\Pr(\hat{s}_u^{\text{IO}} \neq s_{0,u})$ for all $u = 1 \ldots, U$.

Similar to the JO problem, solving IO problem, requires high computational complexity that scales exponentially in the number of UEs [11, 58]. Unlike the JO problem, however, where the computational complexity was due to the searching operation over all possible transmit symbol vectors, the computational complexity for IO problem results in the marginalization in (2.4) [1]. In particular, the summation in (2.4) for the marginalization requires $\mathcal{O}^{U-1}$ operations for each $u$th UE, which scales exponentially in the number of UEs. Thus, solving the IO problem also becomes quickly intractable for a massive MU-MIMO system. In this thesis, we will discuss when our proposed algorithm can achieve the same performance as IO in the asymptotic regime, as well as additional assumptions that we make.

### 2.5.2 Linear Data Detection

As described in Section 2.5.1, solving the IO and JO problem is difficult, especially in massive MU-MIMO systems. Thus, in this section, we discuss linear data detection methods for data detection in MIMO systems. Linear data detection methods are widely used in data detection in massive MU-MIMO because they have, in general, low complexity and can be efficiently mapped into hardware, as

it comprises of additions and multiplications. For the model proposed in Fig. 2.8, we present three different linear data detection methods that provide tradeoffs between complexity and performance.

**Matched Filtering (MF) Data Detection**

Matched filtering, or also known as maximal ratio combining (MRC), is one of the lowest complexity data detection methods, where the estimate is given by:

$$\hat{\mathbf{z}}^{\mathrm{MF}} = \mathrm{diag}(\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y}.$$

MF is shown to be optimal for a fixed number of UEs and infinitely many BS antennas [53], or the low SNR regime [57]. In realistic practical systems, however, MF results in poor spectral efficiency compared to other linear and non-linear data detectors.

**Zero-forcing (ZF) Data Detection**

In a nutshell, ZF data detection "undoes" the effect of the wireless channel by performing a pseudo-inverse of the channel matrix $\mathbf{H}$. The term "zero-forcing" comes from removing the intersymbol interference from other symbols in the noise-free case. ZF problem also amounts to solving for the symbol vector $\tilde{\mathbf{s}}$ that is the closest to the observed output $\mathbf{y}$. The ZF problem can be also obtained from the JO problem in (2.2), where we ignore the prior distribution. Mathematically, ZF problem amounts to solving the following optimization problem. We start with the maximum-likelihood (ML) problem and relax the discrete $\mathcal{O}^U$ set to the

continous complex plane $\mathbb{C}^U$.

$$\begin{aligned}\hat{\mathbf{z}}^{\mathrm{ZF}} &= \min_{\tilde{\mathbf{s}}\in\mathbb{C}^U} \|\mathbf{y}-\mathbf{H}\tilde{\mathbf{s}}\|_2^2 \\ &= \min_{\tilde{\mathbf{s}}\in\mathbb{C}^U} -2\mathrm{Re}\Big(\mathbf{y}^{\mathrm{H}}\mathbf{H}\tilde{\mathbf{s}}\Big) + \|\mathbf{H}\tilde{\mathbf{s}}\|_2^2 \\ &= \min_{\tilde{\mathbf{s}}\in\mathbb{C}^U} -\mathbf{y}^{\mathrm{H}}\mathbf{H}\tilde{\mathbf{s}} - \tilde{\mathbf{s}}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{y} + \tilde{\mathbf{s}}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{H}\tilde{\mathbf{s}}.\end{aligned}$$

Differentiating the above by $\tilde{\mathbf{s}}$ using Wirtinger calculus [60, 61] and equating to zero gives $\hat{\mathbf{z}}^{\mathrm{ZF}} = (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y}$.

Although ZF data detection provides good error-rate performance in the high SNR regime, it performs worse than MF in low SNR regime. In addition, for ill-conditioned channels, ZF can lead to noise enhancement, which also reduces spectral efficiency performance. [57].

**Minimal Mean-Squared Error (MMSE) Data Detection**

The minimum mean-squared error (MMSE) data detector is given as a compromise between MF and ZF data detector. In low and high SNR regime, MMSE achieves the same performance as MF and ZF detector, respectively. We will first derive the MMSE detector for complex-valued transmit signal; we will then derive the MMSE detector for a real-valued transmit signal, e.g., BPSK for a complex-valued system. We note that the majority of the literature do not consider deriving the optimal MMSE estimator for real-valued transmit signals in complex-valued systems [57, 62], when such scenarios can be quite common, i.e., when BPSK or PAM constellation are used.

**MMSE for complex-valued system with complex-valued constellation**

The MMSE detector for a complex-valued transmit signal is obtained by picking the matrix $\mathbf{W}$ that minimizes the mean-squared error (MSE):

$$\begin{aligned}\text{MSE} &= \min_{\mathbf{W}} \mathbb{E}_{\mathbf{s},\mathbf{n}} \left[ \|\mathbf{s} - \mathbf{W}\mathbf{y}\|_2^2 \right] = \min_{\mathbf{W}} \text{tr} \left[ \mathbb{E}_{\mathbf{s},\mathbf{n}} \left[ (\mathbf{s} - \mathbf{W}\mathbf{y})(\mathbf{s}^{\text{H}} - \mathbf{y}^{\text{H}}\mathbf{W}^{\text{H}}) \right] \right] \\ &= \min_{\mathbf{W}} \text{tr} \left[ E_s \mathbf{I}_U - E_s \mathbf{W}\mathbf{H} - E_s \mathbf{H}^{\text{H}}\mathbf{W}^{\text{H}} + \mathbf{W}(E_s \mathbf{H}\mathbf{H}^{\text{H}} + N_0 \mathbf{I}_B)\mathbf{W}^{\text{H}} \right]\end{aligned}$$

Therefore, taking the Wirtinger derivative [60,61] with respective to $\mathbf{W}$ by treating $\mathbf{W}$ and $\mathbf{W}^{\text{H}}$ as separate variables we have that:

$$\frac{\text{dMSE}}{\text{d}\mathbf{W}^{\text{H}}} = -E_s \mathbf{H}^{\text{H}} + \mathbf{W}(E_s \mathbf{H}\mathbf{H}^{\text{H}} + N_0 \mathbf{I}_B),$$

so that the optimal $\mathbf{W}^{\star}$ for $\alpha = N_0/E_s$ corresponds to

$$\begin{aligned}\mathbf{W}^{\star} = \mathbf{H}^{\text{H}} \Big( \mathbf{H}\mathbf{H}^{\text{H}} + \alpha \mathbf{I}_B \Big)^{-1} &\overset{(a)}{=} \mathbf{H}^{\text{H}} \Big( \alpha^{-1} \mathbf{I}_B - \alpha^{-1} \mathbf{H} (\mathbf{H}^{\text{H}}\mathbf{H} + \alpha \mathbf{I}_U)^{-1} \mathbf{H}^{\text{H}} \Big) \\ &= \alpha^{-1} \Big( \mathbf{I}_U - \mathbf{H}^{\text{H}}\mathbf{H} (\mathbf{H}^{\text{H}}\mathbf{H} + \alpha \mathbf{I}_U)^{-1} \Big) \mathbf{H}^{\text{H}} \\ &\overset{(b)}{=} (\mathbf{H}^{\text{H}}\mathbf{H} + \alpha \mathbf{I}_U)^{-1} \mathbf{H}^{\text{H}},\end{aligned}$$

where $(a)$ follows from Woodbury matrix identity, and $(b)$ follows from noting that $\mathbf{I}_U = (\mathbf{H}^{\text{H}}\mathbf{H} + \alpha \mathbf{I}_U)(\mathbf{H}^{\text{H}}\mathbf{H} + \alpha \mathbf{I}_U)^{-1}$. This transformation is useful when $U \leq B$ as it reduces the $B \times B$ matrix inversion to $U \times U$ matrix inversion. Therefore, the linear MMSE estimate for complex-valued transmit signal for MMSE data detector is given by $\hat{\mathbf{z}}^{\text{LMMSE}} = (\mathbf{H}^{\text{H}}\mathbf{H} + N_0/E_s \mathbf{I}_U)^{-1} \mathbf{H}^{\text{H}}\mathbf{y}$.

Note that the linear MMSE detector is obtained from the JO problem in (2.2) if we assume a Gaussian distribution for the prior $p(\mathbf{s})$. To see this, if we let

$p(\mathbf{s}) = (\pi E_s)^{-U/2} \exp(-\frac{1}{E_s}\|\mathbf{s}\|_2^2)$, we have

$$\begin{aligned}\hat{\mathbf{s}}^{\mathrm{JO}} &= \arg\max_{\tilde{\mathbf{s}}\in\mathbb{C}^U}\left(-\frac{\|\mathbf{y}-\mathbf{H}\tilde{\mathbf{s}}\|_2^2}{N_0} + \log p(\tilde{\mathbf{s}})\right) \\ &= \arg\min_{\tilde{\mathbf{s}}\in\mathbb{C}^U}\left(\|\mathbf{y}-\mathbf{H}\tilde{\mathbf{s}}\|_2^2 + \frac{N_0}{E_s}\|\tilde{\mathbf{s}}\|_2^2\right) \\ &= \arg\min_{\tilde{\mathbf{s}}\in\mathbb{C}^U}\left\|\mathbf{y}'-\mathbf{H}'\tilde{\mathbf{s}}\right\|_2^2,\end{aligned}$$

where we defined $\mathbf{y}' = [\mathbf{y}; \mathbf{0}_{U\times 1}]$ and $\mathbf{H}' = [\mathbf{H}; \sqrt{N_0/E_s}\mathbf{I}_U]$ The solution is given by $\hat{\mathbf{s}}^{\mathrm{JO}} = (\mathbf{H}^{\mathrm{H}}\mathbf{H} + N_0/E_s\mathbf{I}_U)^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y}$, which coincides exactly to the linear MMSE estimate.

**MMSE for complex-valued system with real-valued constellation**

The MMSE detector for a real-valued transmit signal for a complex-valued system is different than complex-transmit signal MMSE derived above. The MMSE detector for a real-valued signal can be derived by first decomposing the complex-valued received vector $\mathbf{y} \in \mathbb{C}^B$, channel matrix $\mathbf{H} \in \mathbb{C}^{B\times U}$, and noise vector $\mathbf{n} \in \mathbb{C}^B$ as $\overline{\mathbf{y}} = \left[\mathrm{Re}(\mathbf{y})^{\mathrm{T}}\ \mathrm{Im}(\mathbf{y})^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{2B}$, $\overline{\mathbf{H}} = \left[\mathrm{Re}(\mathbf{H})^{\mathrm{T}}\ \mathrm{Im}(\mathbf{H})^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{2B\times U}$, and $\overline{\mathbf{n}} = \left[\mathrm{Re}(\mathbf{n})^{\mathrm{T}}\ \mathrm{Im}(\mathbf{n})^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{2B}$. With these new quantities, the MMSE detector for real-valued transmit signal is obtained by picking the matrix $\overline{\mathbf{W}} \in \mathbb{R}^{U\times 2B}$ that minimizes the MSE:

$$\begin{aligned}\overline{\mathrm{MSE}} &= \min_{\overline{\mathbf{W}}} \mathbb{E}_{\mathbf{s},\mathbf{n}}\left[\left\|\mathbf{s}-\overline{\mathbf{W}}\overline{\mathbf{y}}\right\|_2^2\right] = \min_{\overline{\mathbf{W}}} \mathrm{tr}\left[\mathbb{E}_{\mathbf{s},\overline{\mathbf{n}}}\left[(\mathbf{s}-\overline{\mathbf{W}}\overline{\mathbf{y}})(\mathbf{s}^{\mathrm{T}}-\overline{\mathbf{y}}^{\mathrm{T}}\overline{\mathbf{W}}^{\mathrm{T}})\right]\right] \\ &= \min_{\overline{\mathbf{W}}} \mathrm{tr}\left[E_s\mathbf{I}_U - E_s\overline{\mathbf{W}}\overline{\mathbf{H}} - E_s\overline{\mathbf{H}}^{\mathrm{T}}\overline{\mathbf{W}}^{\mathrm{T}} + \overline{\mathbf{W}}\left(E_s\overline{\mathbf{H}}\overline{\mathbf{H}}^{\mathrm{T}} + \frac{N_0}{2}\mathbf{I}_B\right)\overline{\mathbf{W}}^{\mathrm{T}}\right]\end{aligned}$$

Differentiating $\overline{\mathrm{MSE}}$ by $\overline{\mathbf{W}}$, we have:

$$\frac{\mathrm{d}\overline{\mathrm{MSE}}}{\mathrm{d}\overline{\mathbf{W}}^{\mathrm{T}}} = -2E_s\overline{\mathbf{H}}^{\mathrm{T}} + 2\overline{\mathbf{W}}\left(E_s\overline{\mathbf{H}}\overline{\mathbf{H}}^{\mathrm{T}} + \frac{N_0}{2}\mathbf{I}_B\right).$$

Therefore, setting this to zero, we have the optimal linear matrix $\overline{\mathbf{W}}^{\star}$ for $\overline{\alpha} = N_0/(2E_s)$ as

$$\overline{\mathbf{W}}^{\star} = \overline{\mathbf{H}}^{\mathrm{T}}\left(\overline{\mathbf{H}}\overline{\mathbf{H}}^{\mathrm{T}} + \overline{\alpha}\mathbf{I}_B\right)^{-1} = (\overline{\mathbf{H}}^{\mathrm{T}}\overline{\mathbf{H}} + \overline{\alpha}\mathbf{I}_U)^{-1}\overline{\mathbf{H}}^{\mathrm{T}},$$

where we utilized Woodbury matrix identity as done in our derivations for complex-valued MMSE transmit signals. Therefore, the linear MMSE data detector for real-valued transmit signal is given by:

$$\begin{aligned}\hat{\mathbf{z}}^{\mathrm{LMMSE}} = \overline{\mathbf{W}}^{\star}\overline{\mathbf{y}} = &\left(\mathrm{Re}(\mathbf{H})^{\mathrm{T}}\,\mathrm{Re}(\mathbf{H}) + \mathrm{Im}(\mathbf{H})^{\mathrm{T}}\,\mathrm{Im}(\mathbf{H}) + \frac{N_0}{2E_s}\mathbf{I}_U\right)^{-1} \\ &\times\left(\mathrm{Re}(\mathbf{H})^{\mathrm{T}}\,\mathrm{Re}(\mathbf{y}) + \mathrm{Im}(\mathbf{H})^{\mathrm{T}}\,\mathrm{Im}(\mathbf{y})\right).\end{aligned}$$

Similar to the previous derivation for MMSE in complex-valued transmit signals, we obtain the same expression if we solve the JO problem when we assume the prior distribution is a real-valued Gaussian with variance $E_s$, i.e., $\mathbf{s} \sim \mathcal{N}(\mathbf{0}_{U\times 1}, E_s\mathbf{I}_U)$.

CHAPTER 3

# INDIVIDUALLY-OPTIMAL DATA DETECTION IN MASSIVE MU-MIMO

## 3.1 Introduction

We consider the problem of recovering the $U$-dimensional data vector $\mathbf{s}_0 \in \mathcal{O}^U$ from the noisy input-output relation $\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}$, by solving the individually-optimal (IO) data detection problem [11, 63]

$$\text{(IO)} \quad s_u^{\text{IO}} = \arg\max_{\tilde{s}_u \in \mathcal{O}} p(\tilde{s}_u|\mathbf{y}), \quad u = 1, 2, \ldots, U,$$

where $p(\tilde{s}_u|\mathbf{y})$ is the probability density function conditioned on observing the receive vector $\mathbf{y} \in \mathbb{C}^B$ and assuming Gaussian noise for the noise vector $\mathbf{n} \in \mathbb{C}^B$. The scalar $s_u^{\text{IO}}$ corresponds to the $u$th IO estimate, $\mathcal{O}$ is a finite constellation set (e.g., PAM, PSK, or QAM), $U$ and $B$ denote the number of transmitters and receivers, respectively, and $\mathbf{H} \in \mathbb{C}^{B \times U}$ represents the (known) multiple-input multiple-output (MIMO) channel matrix.

We develop a computationally efficient algorithm, referred to as LAMA (short for large MIMO approximate message passing), which is able to achieve the error-rate performance of the IO data-detector under certain assumptions on the MIMO channel matrix and the constellation alphabet. We show that in the large system limit, i.e., for $\beta = U/B$ and $U \to \infty$, and for i.i.d. Rayleigh fading MIMO channels, LAMA decouples the noisy MIMO system into a set of independent additive white Gaussian noise (AWGN) channels with equal signal-to-noise ratio (SNR). LAMA is iterative in nature and enables one to compute the noise variance $\sigma_t^2$ of each decoupled AWGN channel in each iteration $t$. This property allows for a precise analysis of the algorithm's performance (in terms of achievable rates

and error rate) and complexity (in terms of the number of LAMA iterations). Furthermore, we can accurately characterize the performance/complexity trade-offs without the need for expensive system simulations.

### 3.1.1 Application Examples

The considered MIMO system model covers a variety of applications, including the following examples.

#### Massive Multi-User (MU) MIMO

Massive MU-MIMO (also known as large-scale or full-dimensional MIMO) will be a key technology to meet the demands for higher spectral efficiency and quality-of-service-in fifth-generation (5G) wireless systems [5, 30, 64]. Massive MU-MIMO relies on hundreds of antennas at the base-station (BS) that serve tens of users simultaneously and in the same frequency band. This technology promises significant gains in terms of spectral efficiency as well as lower operational power consumption compared to that of existing, small-scale MIMO systems [5]. In addition, in the large BS-antenna limit, i.e., where $B \to \infty$ and the total number $U$ of user antennas remains constant, low-complexity data detection and precoding methods (such as the matched filter) turn out to be optimal [53]. However, as demonstrated in [6, 22, 65], practical (finite-dimensional) antenna configurations require more sophisticated data detection algorithms, which entail high computational complexity. The proposed LAMA algorithm enables high-performance and low-complexity data detection in practical massive MU-MIMO systems with higher-order modulation schemes, and allows for an

accurate prediction of the fundamental performance/complexity trade-offs.

### Code-Division Multiple Access (CDMA)

CDMA is a classical transmission technology, in which multiple users simultaneously access a common resource (such as time or frequency) by modulating their individual information signals using spreading sequences [66–70]. A significant portion of the CDMA literature studied the limits (such as the achievable rates for a given modulation scheme) of randomly spread CDMA. In the considered system model, the spreading matrix corresponds to $\mathbf{H}$ with i.i.d. zero-mean Gaussian entries, $U$ denotes the number of users, and the spreading sequences are of length $B$. For common constellations (such as PAM, PSK, or QAM), we provide conditions that depend on the *system ratio* $\beta = U/B$ (also known as the loading factor) in the large-system limit for which LAMA achieves the same error-rate performance of the IO data detector [11]. Our analysis recovers classical results from the CDMA literature [10,11,71] while providing practical means for closely approaching these limits in finite-dimensional systems at low computational complexity.

### Finding Discrete Solutions to Systems of Linear Equations

The considered system model also enables one to study the recovery of integer solutions to the (noisy) system of linear equations $\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}$. For noiseless observations, i.e., $\mathbf{y} = \mathbf{H}\mathbf{s}$, and for the case of $\mathcal{O}$ being (a subset of) the integers, LAMA is able to perfectly recover $\mathbf{s}_0 \in \mathcal{O}^U$ provided that the entries of the system matrix $\mathbf{H}$ are i.i.d. zero-mean Gaussian distributed and the system ratio

$\beta = U/B$ does not exceed a certain *exact recovery threshold (ERT)*. This result is relevant for solving systems of linear Diophantine equations, which finds, for example, use in number theory, cryptography, or closest vector problems in lattices; see [72–75] and the references therein.

### 3.1.2 Relevant Prior Art

Early results on optimal data detection in large MIMO systems reach back to [76] where Verdú and Shamai analyzed the spectral efficiency of multi-user detectors in randomly-spread CDMA systems. The authors provided a precise characterization of the achievable rates with optimal data detection and demonstrated that the system's randomness (due to the random spreading sequences) disappears in the large-system limit. Tanaka [10] derived analytical expressions for the error-rate performance and the multi-user efficiency (equivalent to the noise variance in a single AWGN channel) for the IO data detector using the replica method [77]; Tanaka's results were obtained for BPSK constellations using the replica method in [10] and later proven rigorously in [78]. Guo and Verdú provided an extension of these results to arbitrary discrete inputs [11]. Moreover, it was shown that for a certain family of multi-user detectors, referred as posterior mean estimators (PMEs), the communication system decouples into a set of parallel and independent AWGN channels with equal SNR [10–12, 79]. All of these results study the fundamental performance of IO detection in the large-system limit, i.e., for $\beta = U/B$ with $U \to \infty$. Corresponding practical algorithms have been proposed for BPSK in real-valued systems [71], [80]—in contrast, LAMA is a practical algorithm for general constellations and complex-valued systems, and enables a corresponding theoretical performance analysis.

LAMA builds upon approximate message passing (AMP) [1,81,82], which was initially proposed for sparse signal recovery and compressive sensing [83–85]. In the large-system limit, the estimates obtained by AMP correspond to the true signal perturbed by i.i.d. Gaussian noise [86]. In addition, the variance of the Gaussian random variables in each AMP iteration can be tracked exactly via the state evolution (SE) framework [81, 82]; this feature enables an exact performance analaysis. AMP has been generalized to i.i.d. signal priors using the Bayesian AMP framework [1, 87, 88] and to sparse recovery in complex-valued systems [9]. More recently, AMP and the SE framework have been extended to more general observation models in [59,89,90]. Within the last few years, AMP has been successfully deployed in a variety of applications [91–94], including signal restoration [95,96], imaging [97], phase retrieval [98], and denoising [85,99]. AMP-related algorithms have also been used for data detection in many different communication systems [95, 100–103]. While these results showcase the potential of AMP for data detection in wireless systems, they lack of a rigorous performance analysis. In this chapter of the thesis, we focus on a theoretical performance analysis of AMP for data detection and provide conditions for which it achieves IO performance.

## 3.2 Complex Bayesian Approximate Message Passing (cB-AMP)

Before we delve into the LAMA algorithm and discuss its performance, we first present the complex Bayesian Approximate Message Passing algorithm.

### 3.2.1 System Model and Assumptions

We start by developing the complex Bayesian AMP (cB-AMP) framework, which builds the foundation of the LAMA algorithm, which we will discuss in Section 3.3.2. We specify our model assumptions, derive cB-AMP, and detail the complex-valued state-evolution (cSE) framework.

Before we present the complex Bayesian Approximate Message Passing (cB-AMP) algorithm, we will present additional definitions and assumptions that we will use throughout this section. Unlike classical wireless literature where $U \leq B$ is commonly assumed, we do not impose any restrictions on the so-called *system ratio* (also known as the loading factor in CDMA literature [104]), which we define as $\beta = U/B$. We will often use the following definition:

**Definition 1.** *For a MIMO system with U and B transmitters and receivers respectively, we define the* large-system limit *by fixing the system ratio $\beta = U/B$ and letting $U \to \infty$.*

In what follows, we will consider underdetermined ($\beta \leq 1$) as well as overdetermined ($\beta > 1$) systems. We will frequently use of the following assumptions on the MIMO system matrix $\mathbf{H}$ [86]:

**(A1)** The entries of $\mathbf{H}$ are normalized so that the columns have zero mean and unit $u_2$-norm. In addition, the real and imaginary parts are independent with identical variance. Furthermore, all entries have similar magnitude $O(1/\sqrt{B})$, and all entries of $\mathbf{H}$ are pairwise independent.

**(A2)** The entries of $\mathbf{H}$ are i.i.d. circularly-symmetric complex Gaussian, i.e., $H_{k,u} \sim \mathcal{CN}(0, 1/B)$, $\forall k, u$.

We note that (A2) implies (A1) in the large-system limit; see [86] for the details. Based on the assumption (A1) and (A2), the average received SNRis given by:

$$\mathsf{SNR} = \frac{\mathbb{E}_{\mathbf{H},\mathbf{s}_0}[\|\mathbf{H}\mathbf{s}_0\|_2^2]}{\mathbb{E}_{\mathbf{n}}[\|\mathbf{n}\|_2^2]} = \frac{U}{B} \cdot \frac{E_s}{N_0} = \beta \frac{E_s}{N_0}, \tag{3.1}$$

where the signal power is given by $E_s = \mathbb{E}[|s_{0,u}|^2]$ for all $u = 1, \ldots, U$.

We will also consider the case in which the receiver assumes the following (possibly) mismatched input-output relation:

$$\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}^{\text{post}}. \tag{3.2}$$

Here, $\mathbf{n}^{\text{post}} \sim \mathcal{CN}(\mathbf{0}_{B\times 1}, N_0^{\text{post}}\mathbf{I}_B)$ models noise with postulated noise variance $N_0^{\text{post}}$ (not necessarily equal to $N_0$) and the rest is identical to the assumptions of (2.1). Model (3.2) allows us to analyze a mismatch between the true noise variance $N_0$ and the postulated noise variance $N_0^{\text{post}}$ assumed by the detector. The case $N_0 = N_0^{\text{post}}$ corresponds to an ideal system with perfect knowledge of the noise variance.

To arrive at an efficient algorithm that achieves the same error-rate performance as the IO data detector, we start with the Bayesian AMP (B-AMP) algorithm proposed in [1, 87, 88] to obtain a marginalized distribution $p(\tilde{s}_u|\mathbf{y})$ for each stream $u$. With the marginalized distribution, B-AMP enables the estimation of a vector $\mathbf{s}_0$ from a real-valued version of the system model (2.1). While B-AMP can—in certain cases—be applied to complex-valued systems using the well-known real-valued decomposition,[1] the effective, real-valued system matrix

[1]The complex-valued model (2.1) can be rewritten as the following real-valued model:

$$\begin{bmatrix} \text{Re}(\mathbf{y}) \\ \text{Im}(\mathbf{y}) \end{bmatrix} = \begin{bmatrix} \text{Re}(\mathbf{H}) & -\text{Im}(\mathbf{H}) \\ \text{Im}(\mathbf{H}) & \text{Re}(\mathbf{H}) \end{bmatrix} \begin{bmatrix} \text{Re}(\mathbf{s}) \\ \text{Im}(\mathbf{s}) \end{bmatrix} + \begin{bmatrix} \text{Re}(\mathbf{n}) \\ \text{Im}(\mathbf{n}) \end{bmatrix}.$$

$\overline{\mathbf{H}} \in \mathbb{R}^{2B\times 2U}$ (i) violates the independence assumptions on the entires of $\mathbf{H}$ of (A1), and (ii) inhibits the use of non-separable symbol alphabets, such as phase-shift keying (PSK) constellations. To overcome both of these drawbacks, we next develop a complex-valued version of B-AMP, which we refer to as cB-AMP. We start with Bayes' rule and factorize

$$p(\mathbf{y}|\mathbf{s})p(\mathbf{s}) = \prod_{b=1}^{B} p(y_b|\mathbf{s}) \prod_{u=1}^{U} p(s_u), \tag{3.3}$$

where we assume (i) complex Gaussian noise with postulated noise variance $N_0^{\text{post}}$ given by

$$p(y_b|\mathbf{s}) = \frac{1}{Z} \exp\Big( -\frac{1}{N_0^{\text{post}}} |y_b - \mathbf{h}_b^{\text{row}}\mathbf{s}|^2 \Big),$$

with the constant $Z$ so that $\int_{\mathbb{C}} p(y_b|\mathbf{s})\mathrm{d}y_b = 1$, and (ii) that the transmitted symbols are i.i.d.

To arrive at an efficient inference method, we deploy the sum-product message-passing algorithm [105]. However, as noted in [81], a corresponding full-fledged message passing scheme is impractical. Hence, as in [1,86], we simplify the algorithm by assuming a Gaussian distribution for the marginal densities of the messages $p(\hat{s}_u|s_u) \sim \mathcal{CN}(s_u, \tau^2)$, i.e., $\hat{s}_u = s_u + \tau N_{\mathbb{C}}$, where $N_{\mathbb{C}} \sim \mathcal{CN}(0,1)$, so that [86] the posterior distribution $f(s_u|\hat{s}_u)$ is given by

$$\begin{aligned} f(s_u|\hat{s}_u) &= \frac{p(\hat{s}_u|s_u)p(s_u)}{p(\hat{s}_u)} \\ &= \frac{1}{Z'} \exp\left( -\frac{1}{\tau^2}|s_u - \hat{s}_u|^2 \right) p(s_u), \end{aligned} \tag{3.4}$$

with the normalization constant $Z'$. We denote the conditional mean $\mathsf{F}\big(\hat{s}_u, \tau^2\big)$ and variance $\mathsf{G}\big(\hat{s}_u, \tau^2\big)$ of a random variable $S_u$ distributed according to (3.4) as the message mean and message variance for the relation $\hat{s}_u = S_u + \tau N_{\mathbb{C}}$; both

quantities are defined as follows:

$$\mathsf{F}(\hat{s}_u, \tau^2) = \mathbb{E}_{S_u}[S_u | \hat{s}_u = S_u + \tau N_{\mathbb{C}}], \tag{3.5}$$

$$\mathsf{G}(\hat{s}_u, \tau^2) = \mathbb{V}\mathrm{ar}_{S_u}[S_u | \hat{s}_u = S_u + \tau N_{\mathbb{C}}]. \tag{3.6}$$

By using the methods developed in [1, 9], we can simplify the sum-product message-passing computations for (3.3) which stems from the Gaussian assumption for the marginal densities of the messages. We refer to the resulting algorithm as complex Bayesian AMP (cB-AMP), which is summarized below (and derived in detail in Appendix A.1.2):

**Algorithm 1.** *Suppose that* $\mathbf{H}$ *satisfies (A1) and [1, Lem. 5.56] holds. Then, the* complex Bayesian AMP (cB-AMP) *algorithm performs the following steps for each iteration* $t = 1, 2, \ldots,$*:*

$$\begin{aligned}
\hat{\mathbf{z}}^t &= \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}} \mathbf{r}^t \\
\hat{\gamma}_t^2 &= N_0^{\mathrm{post}} + \hat{\tau}_t^2 \\
\hat{\mathbf{s}}^{t+1} &= \mathsf{F}(\hat{\mathbf{z}}^t, \hat{\gamma}_t^2) \\
\hat{\tau}_{t+1}^2 &= \beta \big\langle \mathsf{G}(\hat{\mathbf{z}}^t, \hat{\gamma}_t^2) \big\rangle \\
\mathbf{r}^{t+1} &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \frac{\beta \mathbf{r}^t}{2} \Big\langle \Big( \partial_1 \mathsf{F}^{\mathrm{R}} + \partial_2 \mathsf{F}^{\mathrm{I}} \Big) (\hat{\mathbf{z}}^t, \hat{\gamma}_t^2) \Big\rangle \\
&\quad - \mathrm{i} \frac{\beta \mathbf{r}^t}{2} \Big\langle \Big( \partial_2 \mathsf{F}^{\mathrm{R}} - \partial_1 \mathsf{F}^{\mathrm{I}} \Big) (\hat{\mathbf{z}}^t, \hat{\gamma}_t^2) \Big\rangle
\end{aligned} \tag{3.7}$$

*where the algorithm is initialized* $\hat{s}_u^1 = \mathbb{E}_{S_u}[S_u]$ *for all* $u = 1, \ldots, U$*,* $\hat{\tau}_1^2 = \beta E_s$ *and* $\mathbf{r}^1 = \mathbf{y} - \mathbf{H}\mathbf{s}^1$*. The functions* $\partial_{\{1,2\}} \mathsf{F}^{\{\mathrm{R,I}\}}(x + iy, \tau^2)$ *are defined as*

$$\begin{aligned}
\partial_1 \mathsf{F}^{\mathrm{R}} &\triangleq \frac{\partial \mathrm{Re}\big(\mathsf{F}(x + \mathrm{i}y, \tau^2)\big)}{\partial x}, \quad \partial_2 \mathsf{F}^{\mathrm{R}} \triangleq \frac{\partial \mathrm{Re}\big(\mathsf{F}(x + \mathrm{i}y, \tau^2)\big)}{\partial y}, \\
\partial_1 \mathsf{F}^{\mathrm{I}} &\triangleq \frac{\partial \mathrm{Im}\big(\mathsf{F}(x + \mathrm{i}y, \tau^2)\big)}{\partial x}, \quad \partial_2 \mathsf{F}^{\mathrm{I}} \triangleq \frac{\partial \mathrm{Im}\big(\mathsf{F}(x + \mathrm{i}y, \tau^2)\big)}{\partial y},
\end{aligned}$$

*and* $\partial_{\{1,2\}} \mathsf{F}^{\{R,I\}}$*,* $\mathsf{F}$*, as well as* $\mathsf{G}$ *operate element-wise on vectors.*

We note that $\hat{\mathbf{s}}^{t+1}$ in Algorithm 1 corresponds to the (nonlinear) minimum mean-squared error (MMSE) estimate defined in (3.5). For a real-valued system with $N_0 = N_0^{\text{post}}$, cB-AMP reduces to the real-valued Bayesian AMP (B-AMP) proposed in [1]; a short proof is given in Appendix A.1.3.

**Lemma 1.** *Let $N_0^{\text{post}} = N_0$ and assume $\mathbf{H}$ satisfies (A1). If $\mathbf{H}$, $\mathbf{s}$, and $\mathbf{n}$ are real-valued, then cB-AMP reduces to B-AMP in [1].*

### 3.2.2 cSE: Complex State Evolution (with Mismatch)

Two unique features of AMP-based algorithms are (i) the output decouples the system into parallel independent channels with additive Gaussian noise (see Fig. 3.1 for an illustration), and (ii) the noise variance of the decoupled AWGN channel can be predicted analytically via fixed-point equations in the large-system limit, which is known as *state evolution* (SE) [86]. The SE framework has been investigated in detail in [1] for B-AMP and in [9] for CAMP, which is a special case of cB-AMP proposed here.[2]

Before we delve into the complex SE (cSE) framework for analysis on the noise variance of the decoupled AWGN channels, we first define the mean-squared error (MSE) of the MMSE output of the cB-AMP algorithm.

**Definition 2.** *Suppose that $\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}$, where the signal $\mathbf{s}_0$ is distributed according to $\mathbf{s}_0 \sim p(\mathbf{s}_0)$, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{B\times 1}, N_0\mathbf{I}_B)$, and the postulated noise variance is $N_0^{\text{post}}$. Let $\hat{\mathbf{s}}^{t+1}$ be the MMSE output of cB-AMP after $t$ iterations. We define the MSE of the*

[2] The SE framework presented [9] focused on sparse signal recovery; we present SE framework for general prior distributions.

*MMSE output of cB-AMP after $t$ iterations as follows:*

$$\mathrm{MSE}_t = \lim_{U\to\infty} \frac{1}{U}\|\hat{\mathbf{s}}^{t+1} - \mathbf{s}_0\|_2^2 = \lim_{U\to\infty} \frac{1}{U}\sum_{u=1}^{U}\left|\mathsf{F}(\hat{z}_u^t, \hat{\gamma}_t^2) - s_{0,u}\right|^2. \tag{3.8}$$

**Definition 3.** *The* effective noise variance *for the MMSE estimate of cB-AMP after $t$ iterations is given by*

$$\sigma_{t+1}^2 = \lim_{B\to\infty} \frac{1}{B}\|\mathbf{r}^{t+1}\|_2^2 = N_0 + \beta MSE_t. \tag{3.9}$$

We note that a proof of (3.9) was given in [106, Lem. 4.1]. While $\sigma_{t+1}^2$ corresponds to the effective noise variance (illustrated in Fig. 3.1(b)), the postulated output variance $\gamma_{t+1}^2$ defined below corresponds to the *predicted* value of $\sigma_{t+1}^2$ at iteration $t$ of cB-AMP. If there is a mismatch in the noise variance $N_0^{\text{post}} \neq N_0$, then the postulated output variance $\gamma_t^2$ differs from the actual noise variance $\sigma_t^2$, i.e., $\gamma_t^2 \neq \sigma_t^2$.

**Definition 4.** *The* postulated output variance *of cB-AMP after $t$ iterations is given by*

$$\gamma_{t+1}^2 = \lim_{U\to\infty} \hat{\gamma}_{t+1}^2 = N_0^{\text{post}} + \beta \lim_{U\to\infty} \frac{1}{U}\sum_{u=1}^{U} \mathsf{G}(\hat{z}_u^t, \hat{\gamma}_t^2). \tag{3.10}$$

Using the definitions of the effective noise variance and postulated output variance, we can formulate the complex SE (cSE) framework with noise variance mismatch for cB-AMP. The complex SE framework was proven rigorously in [82].

**Theorem 2.** *Suppose the entries of $\mathbf{s}_0$ are i.i.d. $p(\mathbf{s}_0) \sim \prod_{u=1}^{U} p(s_{0,u})$ and the entries of the MIMO system matrix $\mathbf{H}$ satisfy (A2). Let $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{B\times 1}, N_0\mathbf{I}_B)$ and $\mathsf{F} : \mathbb{C} \to \mathbb{C}$ be a pseudo-Lipschitz function as defined in [82, Sec. 1.1, Eq. 1.5]. Assume the large-system limit and that the postulated noise variance is $N_0^{\text{post}}$. Then, the effective noise variance $\sigma_{t+1}^2$ in (3.9) and postulated output variance $\gamma_{t+1}^2$ in (3.10) of cB-AMP in iteration $t$ are*

*given by the following coupled recursion:*

$$\sigma_{t+1}^2 = N_0 + \beta \Psi(\sigma_t^2, \gamma_t^2), \tag{3.11}$$

$$\gamma_{t+1}^2 = N_0^{\text{post}} + \beta \Phi(\sigma_t^2, \gamma_t^2). \tag{3.12}$$

*The MSE function* $\Psi$ *and variance function* $\Phi$ *are defined by*

$$\Psi(\sigma_t^2, \gamma_t^2) = \mathbb{E}_{S, N_{\mathbb{C}}} \left[ \left| \mathsf{F}(S + \sigma_t N_{\mathbb{C}}, \gamma_t^2) - S \right|^2 \right], \tag{3.13}$$

$$\Phi(\sigma_t^2, \gamma_t^2) = \mathbb{E}_{S, N_{\mathbb{C}}} \left[ \mathsf{G}(S + \sigma_t N_{\mathbb{C}}, \gamma_t^2) \right], \tag{3.14}$$

*respectively, with* $S \sim p(S)$*. The recursion is initialized at* $t = 1$ *with*

$$\sigma_1^2 = N_0 + \beta \, \mathbb{V}\mathrm{ar}_S[S] \quad \textit{and} \quad \gamma_1^2 = N_0^{\text{post}} + \beta \, \mathbb{V}\mathrm{ar}_S[S].$$

We note that the MSE function $\Psi(\sigma^2, \sigma^2)$ is identical to the "$\mathsf{mmse}(\mathsf{snr})$" function in [11, 104, 107, 108] with the relation $\mathsf{snr} = 1/\sigma^2$, used to derive the relationship between the mutual information and the MSE function. We also note that the MSE of cB-AMP at iteration $t$ as defined in (3.8) is equivalent to $\Psi(\sigma_t^2, \gamma_t^2)$ in the large-system limit. Theorem 2 implies that the effective noise variance of cB-AMP $\sigma^2$ can, in the large-system limit, be predicted exactly by the variance of a *single* random variable mixed with additive Gaussian noise. If $N_0 = N_0^{\text{post}}$, then we arrive at the following result.

**Corollary 3.** *Let* $N_0^{\text{post}} = N_0$ *in Theorem 2. Then (3.11) is identical to (3.12), and the cSE reduces to the following recursion:*

$$\sigma_{t+1}^2 = N_0 + \beta \Psi(\sigma_t^2, \sigma_t^2). \tag{3.15}$$

The proof of Corollary 3 follows from the fact that the MSE equals to the conditional variance, i.e., $\Phi(\sigma_t^2, \sigma_t^2) = \Psi(\sigma_t^2, \sigma_t^2)$. We note that Corollary 3 corresponds to the cSE derived originally in [9] in absence of noise-variance mismatch.

Furthermore, for real-valued systems, Corollary 3 coincides with the original SE framework in [1,81].

## 3.3 LAMA: Large MIMO Approximate Message Passing

Now we have derived the complex Bayesian AMP algorithm, we now derive the LAMA algorithm. We specify the missing aspects of the large-MIMO system model and detail the LAMA algorithm along with the corresponding cSE framework.

### 3.3.1 Large MIMO and Optimal Data Detection

We consider a communication system in which the entries $s_u$, $u = 1, \ldots, U$, of the transmit data vector $\mathbf{s}$ are taken from a finite constellation set $\mathcal{O} = \{a_o : o = 1, \ldots, |\mathcal{O}|\}$ with points $a_o$ chosen, from e.g., a pulse amplitude modulation (PAM), phase-shift keying (PSK), or quadrature amplitude modulation (QAM) alphabet. We assume i.i.d. priors $p(\mathbf{s}) = \prod_{u=1}^{U} p(s_u)$, with the following distribution for each transmit symbol $s_u$:

$$p(s_u) = \sum_{a \in \mathcal{O}} p_a \delta(s_u - a). \tag{3.16}$$

Here, $p_a$ is the (known) prior probability of each constellation point $a \in \mathcal{O}$ with $\sum_{a \in \mathcal{O}} p_a = 1$ and $\delta(\cdot)$ is the Dirac delta distribution; for uniform priors we have $p_a = 1/|\mathcal{O}|$.

The vector $\mathbf{s}_0$ is transmitted through a MIMO channel as in (2.1). We assume perfect knowledge of the MIMO system matrix $\mathbf{H}$ at the receiver and the noise

vector $\mathbf{n}$ to be i.i.d. circularly complex Gaussian with variance $N_0$ per complex entry. As discussed in Section 2.5.1, the individually optimal (IO) data-detection problem in [11, 63] is given by (2.3)

$$\hat{s}_u^{\mathrm{IO}} = \arg\max_{\tilde{s}_u \in \mathcal{O}} \sum_{\tilde{\mathbf{s}}_u \in \mathcal{O}_u^{(U-1)}} \exp\left( -\frac{\|\mathbf{y} - \mathbf{H}\tilde{\mathbf{s}}\|^2}{N_0} + \log p(\tilde{\mathbf{s}}) \right), \tag{3.17}$$

where $\mathcal{O}_u^{(U-1)}$ stands for the subset of $\mathcal{O}^U$ that excludes the $u$th entry and $\tilde{\mathbf{s}}_u \in \mathcal{O}_u^{(U-1)}$ is a $U-1$ dimensional vector from this subset. As mentioned in Section 2.5.1, the IO problem in (3.17) is of combinatorial nature and requires prohibitive complexity in systems with large $U$ [58, 109, 110]. The main complexity for the IO or element-wise MAP problem, in contrast to (2.2), is that the complexity is due to the summation operator across $|\mathcal{O}|^{U-1}$ symbols. We note that the IO data detection achieves the minimum probability of symbol errors (see [58, Sec. 4.1] for a detailed discussion). While computationally efficient algorithms exist for small-scale MIMO systems (up to about eight transmit streams), such as sphere-decoding (SD) based methods [2–4], their average computational complexity still scales exponentially in $U$ [109, 110].[3] Consequently, such methods are not suitable for large MIMO systems. In order to enable data detection for such systems, a variety of sub-optimal algorithms have been proposed in the past; see, e.g., [7, 22, 28, 29, 111, 112] and the references therein.

Instead of solving the IO problem in (3.17) directly, we first compute the marginalized distribution $p(s_u|\mathbf{y})$ using cB-AMP as in Algorithm 1. Once we obtain the marginalized distribution $p(s_u|\mathbf{y})$, $u = 1, \ldots, U$, the IO data-detection problem is transformed in an entry-wise data detection problem that can be solved at low complexity.

[3] In the case of BPSK transmission, soft-input soft-output MAP detectors, such as the one in [4], can exactly solve the IO problem at low average computational complexity for a small number of transmit streams $U$. For higher-order modulation schemes, no known method exists to solve (3.17) at low complexity.

### 3.3.2 Derivation of the LAMA Algorithm

We now use cB-AMP together with the prior $p(\mathbf{s})$ in (3.16) to develop LAMA (short for large MIMO approximate message passing), which enables us to obtain the same error-rate performance of IO data detection at low computational complexity, given certain assumptions on the MIMO system hold (see Section 3.4.3 for precise optimality conditions).

With the prior distribution in (3.16), we can write the posterior distribution (3.4) for the transmit symbol $s_u$ as

$$f(s_u|\hat{s}_u) = \frac{1}{Z}\exp\left(-\frac{|s_u-\hat{s}_u|^2}{\tau^2}\right)\sum_{a\in\mathcal{O}} p_a\delta(s_u-a), \tag{3.18}$$

where we defined a normalization constant $Z$ so that $\int_{\mathbb{C}} f(s_u|\hat{s}_u)\mathrm{d}s_u = 1$, which is given by:

$$Z = \sum_{a\in\mathcal{O}} p_a \exp\left(-\frac{1}{\tau^2}|\hat{s}_u-a|^2\right).$$

With the posterior distribution defined in (3.18), we proceed to computing the message mean and variance. First, the message mean in (3.5) is computed below in (3.19) as follows:

$$\begin{aligned}\mathsf{F}(\hat{s}_u,\tau^2) = \int_{\mathbb{C}} s_u f(s_u|\hat{s}_u,\tau)\,\mathrm{d}s_u &= \frac{\sum_{a\in\mathcal{O}} a p_a \exp\left(-\frac{1}{\tau^2}|\hat{s}_u-a|^2\right)}{\sum_{a'\in\mathcal{O}} p_{a'}\exp\left(-\frac{1}{\tau^2}|\hat{s}_u-a'|^2\right)} \\ &= \sum_{a\in\mathcal{O}} w_a(\hat{s}_u,\tau^2)a,\end{aligned} \tag{3.19}$$

where we use the shorthand notation

$$w_a(\hat{s}_u,\tau^2) = \frac{p_a\exp\left(-\frac{1}{\tau^2}|\hat{s}_u-a|^2\right)}{\sum_{b\in\mathcal{O}} p_b\exp\left(-\frac{1}{\tau^2}|\hat{s}_u-b|^2\right)}.$$

Second, the message variance in (3.6) is computed below in (3.20) as follows:

$$\mathsf{G}(\hat{s}_u, \tau^2) = \int_{\mathbb{C}} |s_u|^2 f(s_u|\hat{s}_u)\, \mathrm{d}s_u - \left|\mathsf{F}(\hat{s}_u, \tau^2)\right|^2,$$

which can be simplified to

$$\mathsf{G}(\hat{s}_u, \tau^2) = \sum_{a \in \mathcal{O}} w_a(\hat{s}_u, \tau^2) \left| a - \mathsf{F}(\hat{s}_u, \tau^2) \right|^2. \tag{3.20}$$

The final step in the derivation of LAMA involves a simplification of the partial derivatives of (3.7) in Algorithm 1. The result is summarized by Lemma 4 with proof given in Appendix A.1.5.

**Lemma 4** (Message variance of the LAMA algorithm)**.** *Suppose that the assumptions of Algorithm 1 hold, and the mean* $\mathsf{F}(\hat{s}_u, \tau^2)$ *as well as the variance* $\mathsf{G}(\hat{s}_u, \tau^2)$ *functions are given by (3.19) and (3.20), respectively. Then, the message variance is given by:*

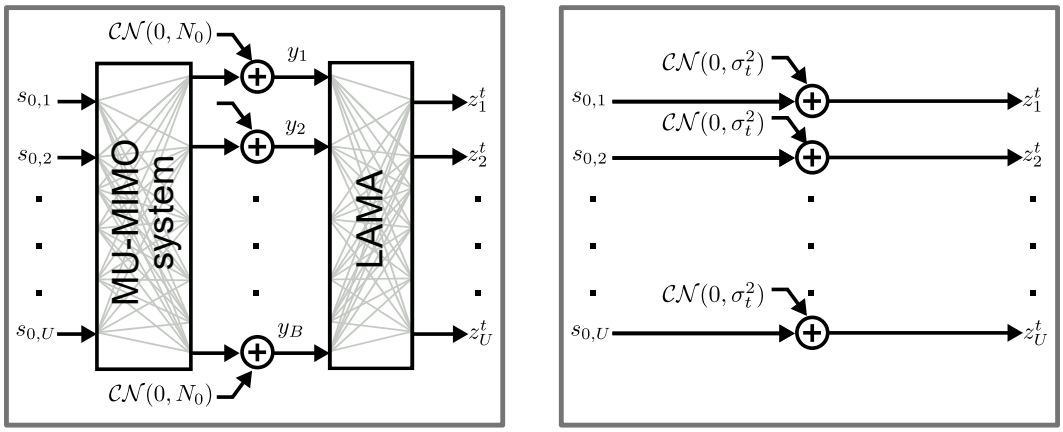
$$\mathsf{G}(\hat{s}_u, \tau^2) = \frac{\tau^2}{2} \left[ \partial_1 \mathsf{F}^R + \partial_2 \mathsf{F}^I \right] \left( \hat{s}_u, \tau^2 \right)$$

*and cB-AMP leads to Algorithm 2.*

With Lemma 4 and Algorithm 1, we arrive at the LAMA algorithm summarized next.

**Algorithm 2** (LAMA)**.** *Suppose that* $\mathbf{H}$ *satisfies (A1) and [1, Lem. 5.56] holds. Then, the* LAMA *algorithm is given by following procedure*

$$\begin{aligned}
\hat{\mathbf{z}}^t &= \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}} \mathbf{r}^t \\
\hat{\gamma}_t^2 &= N_0^{\mathrm{post}} + \hat{\tau}_t^2 \\
\hat{\mathbf{s}}^{t+1} &= \mathsf{F}(\hat{\mathbf{z}}^t, \hat{\gamma}_t^2) && (3.21) \\
\hat{\tau}_{t+1}^2 &= \beta \langle \mathsf{G}(\hat{\mathbf{z}}^t, \hat{\gamma}_t^2) \rangle && (3.22) \\
\mathbf{r}^{t+1} &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \frac{\hat{\tau}_{t+1}^2}{\hat{\gamma}_t^2} \mathbf{r}^t && (3.23)
\end{aligned}$$




(a) MU-MIMO system with LAMA (b) Equivalent decoupled system

Figure 3.1: LAMA decouples an i.i.d. Rayleigh fading large MU-MIMO system (a) into a set of parallel and independent AWGN channels with equal noise variance (b) in the large-system limit.

*for each iteration* $t = 1, 2, \ldots$. *The LAMA algorithm is initialized at iteration* $t = 1$ *with* $\hat{\mathbf{s}}^t = \mathbb{E}_S[S]\mathbf{1}_{U\times 1}$, *where* $S \sim p(S)$, $\mathbf{r}^t = \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^t$, *and* $\hat{\tau}_t^2 = \beta\, \mathbb{V}\text{ar}[S]$.

The main difference between the cB-AMP in Algorithm 1 and LAMA in Algorithm 2 is that the update in (3.7) for cB-AMP is simplified to (3.23) for LAMA and we utilize the prior distribution $p(S)$ to initialize the algorithm. We note that LAMA as summarized in Algorithm 2 makes use of the postulated noise variance $N_0^{\text{post}}$; this allows us not only to model a mismatch in the noise variance, but also enables us to perform IO detection and matched filter (MF) data detection solely by selecting appropriate values for $N_0^{\text{post}}$; see Section 3.3.4.

### 3.3.3 LAMA Decouples Large-MIMO Systems

We now show that LAMA decouples a MIMO system into a set of parallel and independent AWGN channels with identical noise variance in the large system limit (cf. Figs. 3.1(a) and 3.1(b)). First, we discuss the outputs of LAMA: (i) the

Gaussian output vector $\mathbf{z}^t$, (ii) the postulated variance $\hat{\gamma}_t^2$, and (iii) the non-linear MMSE output vector $\hat{\mathbf{s}}^t$.

**(i) Gaussian output vector $\mathbf{z}^t$**

In each iteration $t$, cB-AMP computes the marginal distribution for $s_u$ for $u = 1, \ldots, U$, which corresponds to a Gaussian distribution centered around the original signal $s_{0,u}$ with variance $\sigma_{t+1}^2$. These properties on $\hat{\mathbf{z}}^t$ follow from Theorem 2, which shows that $\hat{\mathbf{z}}^t = \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t$ is distributed according to $\mathcal{CN}(\mathbf{s}_0, \sigma_t^2 \mathbf{I}_U)$ in the large-system limit [82, 106]. Therefore, the input–output relation for each transmit stream $\hat{z}_u^t = \hat{s}_u^t + (\mathbf{h}_u^{\mathrm{col}})^{\mathrm{H}}\mathbf{r}^t$ is equivalent to the following single-input single-output AWGN channel:

$$\hat{z}_u^t = s_{0,u} + n_u^t. \tag{3.24}$$

Here, $s_{0,u}$ is the transmitted signal for $u$th UE and $n_u^t \sim \mathcal{CN}(0, \sigma_t^2)$ is the AWGN with effective noise variance $\sigma_t^2$ per complex entry. Since $p(\hat{z}_u^t | s_{0,u}) \sim \mathcal{CN}(s_{0,u}, \sigma_t^2)$, the posterior distribution of (3.24) is given by $f(s_{0,u} | \hat{z}_u^t)$ with noise variance $\sigma_t^2$. An immediate consequence of these properties is the fact that LAMA decouples the MIMO system (cf. Fig. 3.1(b)). We note that the decoupling behavior of LAMA was observed for posterior mean estimators (PMEs) in randomly spread CDMA systems [11, 58] for which no practical data detection algorithm was given.

**(ii) Postulated output variance $\hat{\gamma}_t^2$**

In the large-system limit, there exist two noise variances: effective noise variance $\sigma_t^2$ from Definition 3 and the postulated output variance $\gamma_t^2$ from Definition 4.

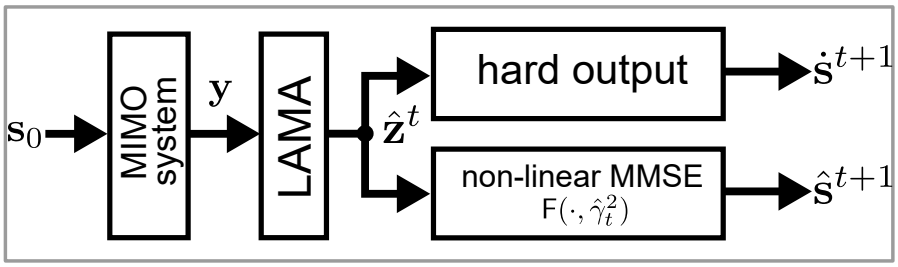
We note that $N_0^{\text{post}} + \hat{\tau}_t^2 = \hat{\gamma}_t^2 \to \gamma_t^2$ in the large-system limit. We clarify the difference between the two quantities below.

The effective noise variance $\sigma_t^2$ is the *true* noise variance in (3.24), whereas the postulated output variance $\gamma_t^2$ is the estimate for $\sigma_t^2$ each iteration $t$. The postulated output variance $\gamma_t^2$ is used as an input to the posterior mean function $\mathsf{F}$ (see (3.21) and Figs. 3.2(a) and 3.2(b)) to the Gaussian vector $\hat{\mathbf{z}}^t$ to obtain the non-linear MMSE estimate $\hat{\mathbf{s}}^{t+1}$. Therefore, when the exact value of $\sigma_t^2$ is unknown at the receiver, a possible performance mismatch can result in using an incorrect value $\gamma_t^2$ for obtaining the MMSE estimate. The cSE framework shown in Theorem 2 enables us to analyze the performance loss due to such a (possible) mismatch in the noise variance $N_0^{\text{post}}$ *exactly*.

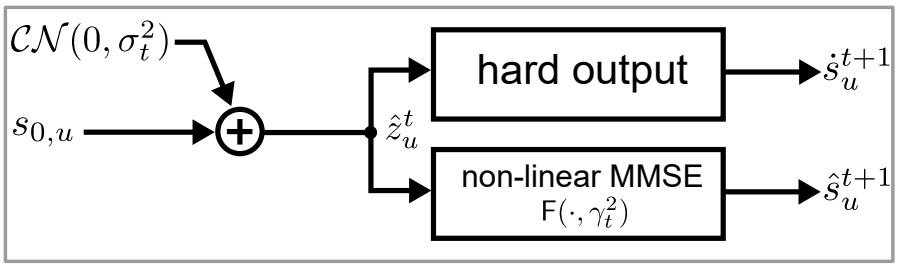
If there is no mismatch in the postulated noise variance, we have $\sigma_t^2 = \gamma_t^2$ by Corollary 3 and hence, the correct noise variance statistic is used for the MMSE estimate in (3.21) every iteration. However, if $N_0^{\text{post}} \neq N_0$, then $\sigma_t^2 \neq \gamma_t^2$, and therefore, LAMA applies the MMSE estimate on the Gaussian vector $\hat{\mathbf{z}}^t$ according to an incorrect statistic, which may cause LAMA to converge to an incorrect solution. To illustrate how LAMA may converge to an incorrect solution, consider the case where $N_0 = 0$, and $N_0^{\text{post}} \to \infty$. In this case, $\hat{\mathbf{z}}^t$ corresponds to the MF detector; see Section 3.3.4 for more details.

**(iii) Non-linear MMSE output vector $\hat{\mathbf{s}}^{t+1}$**

The non-linear MMSE output vector $\hat{\mathbf{s}}^{t+1}$ is given by $\hat{\mathbf{s}}^{t+1} = \mathsf{F}(\mathbf{z}^t, \hat{\gamma}_t^2)$ in (3.19), which can be seen as a conditional mean of the Gaussian output vector $\hat{\mathbf{s}}^t$ for the postulated output variance $\hat{\gamma}_t^2$. The non-linear MMSE output vector $\hat{\mathbf{s}}^{t+1}$




(a) MIMO system with LAMA




(b) Equivalent decoupled system

Figure 3.2: The system with LAMA and its outputs (a) and the statistically-equivalent decoupled AWGN system as seen at the output of LAMA (b). LAMA generates a Gaussian output $\hat{\mathbf{z}}^t$ and a non-linear MMSE estimator output $\hat{\mathbf{s}}^{t+1}$. Hard-output estimates $\dot{s}^{t+1}$ are generated via (3.25). In the large system limit, LAMA decouples the MIMO system into independent, parallel AWGN channels with equivalent output noise variance $\sigma_t^2$.

is identical to the PME [11], where each $u$th output of PME is obtained by the expectation with respect to the conditional distribution $f(s_u|z_u^t)$ with variance $\hat{\gamma}_t^2$ in (3.4).[4] The equivalence of LAMA and the equivalent AWGN relation for the non-linear MMSE estimate is shown in Fig. 3.2. In Fig. 3.2(a), the quantity $\hat{\mathbf{s}}^{t+1}$ is the non-linear MMSE estimate with the postulated noise variance $\hat{\gamma}_t^2$. In the large system limit, the input-output relation for each stream $u$ is an AWGN channel in Fig. 3.2(b) with equivalent variance $\sigma_t^2$ and the postulated variance $\gamma_t^2$.

[4]The conditional distribution $f(s_u \mid z_u^t)$ is called "retro-channel" in [11].

### 3.3.4 LAMA and MF Data Detection

Since LAMA decouples the MIMO system, data detection reduces to element-wise hard decisions for each entry in $\hat{\mathbf{z}}^t$ subject to the postulated output variance $\hat{\gamma}_t^2$ as:

$$\begin{aligned} \dot{s}_u^t &= \arg\max_{\tilde{s}_u \in \mathbb{C}} f(\tilde{s}_u \mid \hat{z}_u^t) \\ &\overset{(a)}{=} \arg\max_{\tilde{s}_u \in \mathbb{C}} \sum_{a \in \mathcal{O}} p_a \delta(\tilde{s}_u - a) \exp\left( -\frac{1}{\hat{\gamma}_t^2} |\tilde{s}_u - \hat{z}_u|^2 \right) \\ &\overset{(b)}{=} \arg\min_{a \in \mathcal{O}} \left\{ \frac{1}{\hat{\gamma}_t^2} |a - \hat{z}_u|^2 - \log p_a \right\}, \end{aligned} \tag{3.25}$$

where in $(a)$, we substituted the posterior distribution (3.18) and in $(b)$, we note that $\exp(\cdot)$ function is monotone.

In this section, we delve into data detection in a greater detail. In particular, we show that by varying the postulated noise variance $N_0^{\text{post}}$ in $\hat{\gamma}_t^2$, LAMA can either perform IO and MF data detection. In particular, (i) for $N_0^{\text{post}} = N_0$, LAMA corresponds to the IO detector and (ii) for $N_0^{\text{post}} \to \infty$, LAMA corresponds to the MF detector. These two "operation modes" are detailed next.

**(i) IO data detection:** $N_0^{\text{post}} = N_0$

Consider $N_0^{\text{post}} = N_0$. From Corollary 3, we have that the equivalent output noise variance and the postulated noise variance are equal, which implies $\sigma_t^2 = \gamma_t^2$ in the large-system limit. Since there is no noise variance mismatch, the output (3.25) achieves the same error-rate performance as the IO data detector which in (2.3) given certain conditions are met; see Section 3.4.5 for precise optimality conditions.

**(ii) MF data detection:** $N_0^{\text{post}} \to \infty$

By letting $N_0^{\text{post}} \to \infty$, it was shown in [11, Eq. (12)] that the output of the non-linear MMSE estimator (3.19) corresponds to the MF output for real-valued signals with $\mathbb{E}[S] = 0$. We now provide conditions for which LAMA with $N_0^{\text{post}} \to \infty$ performs MF data detection for arbitrary system ratios $\beta$. The proof of the following Lemma is given in Appendix A.1.6.

**Lemma 5.** *Fix the constellation set $\mathcal{O}$, and let $S \sim p(S)$. If $\mathbb{E}_S[S] = 0$, then as $N_0^{\text{post}} \to \infty$, the Gaussian output at every iteration $t = 1, 2, \ldots,$ from LAMA corresponds to the MF output:*

$$\lim_{N_0^{\text{post}} \to \infty} \hat{\mathbf{z}}^t = \mathbf{H}^{\text{H}} \mathbf{y}.$$

*If $\mathbb{E}_S[\,\text{Re}(S)\,\text{Im}(S)\,] = \mathbb{E}_S[S|S|^2] = 0$, then, as $N_0^{\text{post}} \to \infty$, the scaled version of the non-linear MMSE estimate also corresponds to the MF output:*

$$\lim_{N_0^{\text{post}} \to \infty} \frac{N_0^{\text{post}}}{E_s} \hat{\mathbf{s}}^t = \mathbf{H}^{\text{H}} \mathbf{y}.$$

Now that we have discussed the "operational modes" of LAMA, we observe the performance of LAMA in the Massive MU-MIMO limit, i.e., $\beta \to 0$.

### 3.3.5 LAMA in the Massive MU-MIMO Limit

We now study the properties of LAMA in the massive MU-MIMO limit, where we fix the number of streams (or layers) $U$ and let the number of BS antennas $B \to \infty$; this corresponds to the case when the system ratio $\beta = U/B \to 0$. As shown in [5, 64], MF data detection is optimal in such scenarios. The following Lemma

reveals that LAMA corresponds to the MF detector in the massive MU-MIMO limit; a proof is given in Appendix A.1.7.

**Lemma 6.** *Assume that $\mathcal{O}$ is fixed and let $N_0^{\text{post}} \geq 0$. Then, for $\beta \to 0$, the Gaussian output $\hat{\mathbf{z}}^t$ of LAMA corresponds to the MF data detector, i.e., $\hat{\mathbf{z}}^t = \mathbf{H}^{\text{H}}\mathbf{y}$, for all $t \geq 1$. Furthermore, the effective noise variance is $\sigma_t^2 = N_0$ for all $t \geq 1$.*

This result is in accordance with [5, 64] and implies that a simple one-shot algorithm (performing a single iteration) is sufficient to perform IO data detection in the massive MU-MIMO limit. Furthermore, LAMA decouples the MIMO system into parallel and independent AWGN channels with variance $\sigma_t^2 = N_0$ (see Fig. 3.1(a) and Fig. 3.1(b)) in every iteration. We emphasize that LAMA can be used in more-realistic massive MU-MIMO systems, i.e., where the number of BS antennas is finite. As we will show in Section 3.5.2, LAMA quickly converges and provides near-optimal performance for realistic massive MU-MIMO antenna configurations.

## 3.4 Optimality of LAMA

Now that we discussed the LAMA algorithm and its properties, we now discuss optimality of LAMA. We now provide exact conditions for which LAMA achieves the performance of the IO data detector. We furthermore study the noiseless case in which LAMA is able to perform error-free data recovery. We also identify the optimality regimes of LAMA in the presence of noise. We start by reviewing existing results on IO and multiuser detection.

### 3.4.1 Existing Results of IO and Multiuser Detection in Large MIMO Systems

An spectral efficiency analysis of IO data detection in large systems with BPSK was presented by Tanaka in [10]. These results were generalized to arbitrary constellation sets in [11]. Under the assumption that replica method is correct, Guo and Verdú showed in [11] that by using PMEs, the multi-user channel in the large-system limit decouples into an AWGN channel for each transmit stream, where the noise is amplified by a factor $\eta^{-1} > 1$ due to the interference of other streams. The factor $\eta \in (0, 1)$, known as the multi-user efficiency, can be computed exactly by solving the following coupled equations for $\eta$ and $\xi$:

$$N_0/\eta = N_0 + \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\left[\left|\mathsf{F}\left(S + \sqrt{N_0/\eta}N_{\mathbb{C}}, N_0/\xi\right) - S\right|^2\right], \tag{3.26}$$

$$N_0/\xi = N_0^{\text{post}} + \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\left[\mathsf{G}\left(S + \sqrt{N_0/\eta}N_{\mathbb{C}}, N_0/\xi\right)\right]. \tag{3.27}$$

Here, the functions $\mathsf{F}$ and $\mathsf{G}$ depend on constellation set $\mathcal{O}$ as in (3.5) and (3.6), respectively.

We note that the performance of IO data detection corresponds to the case with $N_0 = N_0^{\text{post}}$. In this case, the right-hand side of (3.27) is equal to (3.27), and therefore, $\eta = \xi$. Thus, the multi-user efficiency $\eta$ is given by a single fixed-point equation

$$N_0/\eta = N_0 + \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\left[\left|\mathsf{F}\left(S + \sqrt{N_0/\eta}N_{\mathbb{C}}, N_0/\eta\right) - S\right|^2\right]. \tag{3.28}$$

If there exist multiple fixed points to (3.26) and (3.27), we pick the tuple $(\eta, \xi)$

that minimizes the so-called "free energy" (as done in [11, Sec. 2-D]) given by:

$$\begin{aligned}\mathcal{F} = &\int_{\mathbb{C}} q(z, N_0/\eta) \log_2 q(z, N_0/\xi) \mathrm{d}z \\ &+ \frac{1}{\beta}((\xi - 1)\log_2 e - \log_2 \xi) + \log_2 \frac{\xi}{\pi} - \frac{\xi}{\eta}\log_2 e \\ &+ \frac{1}{\beta}\frac{N_0^{\text{post}}}{N_0}\frac{\xi}{\eta}(\eta - \xi)\log_2 e + \frac{1}{\beta}\log_2(2\pi) + \frac{\xi}{\eta\beta}\log_2 e, \end{aligned} \tag{3.29}$$

where the term $q(z, N_0/\eta)$ in (3.29) is obtained by marginalizing the distribution $p(z|s) \sim \mathcal{CN}(s, N_0/\eta)$ with respect to the prior distribution $s \sim p(s)$, i.e., $q(z, N_0/\eta) = \int_{\mathbb{C}} p(z|s)p(s)\mathrm{d}s$.

We note that the aforementioned results rely on the replica method, which build on the replica assumptions [10, 11]. Montanari and Tse in [78] proposed an alternative approach to prove Tanaka's results in [10] up to certain system ratios $\beta$ for BPSK systems. Instead of directly analyzing a dense MIMO system matrix, Montanari and Tse first introduce a "sparse signature" scheme, in which only a sparse subset of the channel matrix is active. For this system, the performance of belief propagation (BP) can be analyzed via density evolution. Once the density evolution expressions were established in the large-system limit, one can "densify" the MIMO system matrix to ensure that the each entry is distributed (A1); we shall refer to this setup as *large-sparse* limit [113]. By doing so, one recovers Tanaka's results derived under the replica method without relying on the replica assumptions. The analysis of BPSK systems using this sparse signature scheme has been generalized to arbitrary prior input distributions in [12, 79, 113]. Not surprisingly, these results agree with the replica results [11] when the fixed-point $\eta$ to (3.28) is unique. In addition, in [113], Wang and Guo showed that BP is equivalent to element-wise MAP estimation, and the detection performance of the BP is identical to that given by a AWGN system with noise amplified by $\eta^{-1}$ obtained in (3.28).

### 3.4.2 Fixed Points of LAMA

Before we provide exact optimality conditions for LAMA, we highlight that under Theorem 2, as $t \to \infty$ the cSE converges to the following fixed-point equations: for $N_0 = N_0^{\text{post}}$, we have

$$\sigma_{\text{IO}}^2 = N_0 + \beta\Psi(\sigma_{\text{IO}}^2, \sigma_{\text{IO}}^2), \tag{3.30}$$

whereas for $N_0^{\text{post}} \neq N_0$, we have

$$\sigma_{\text{m}}^2 = N_0 + \beta\Psi(\sigma_{\text{m}}^2, \gamma_{\text{m}}^2) \quad \text{and} \quad \gamma_{\text{m}}^2 = N_0^{\text{post}} + \beta\Phi(\sigma_{\text{m}}^2, \gamma_{\text{m}}^2). \tag{3.31}$$

As mentioned above, the fixed-point equation for LAMA in (3.30) and (3.31) corresponds to the fixed-point equations for IO data detection in (3.28), and (3.26) and (3.27), respectively, with $\sigma_{\text{m}}^2 = N_0/\eta$ and $\gamma_{\text{m}}^2 = N_0/\xi$.

In general, the above fixed-point equations may have multiple solutions. In the case of a unique fixed point, then LAMA *always* recovers the solution with the minimal effective noise variance $\sigma^2$ regardless of initialization, and thus, achieves the same error-rate performance as IO data detection (see Section 3.4.3 for the details). In the case of such non-unique fixed points, Guo and Verdú choose the solution that minimizes free-energy[5] given in (3.29), whereas the fixed point obtained by LAMA depends on the initialization[6] of the algorithm and thus, we cannot expect it to converge to the same fixed point that minimizes the free-energy (3.29). We note that depending on the initialization of LAMA presented in Algorithm 2, LAMA converges to the fixed-point solution with the largest effective noise variance $\sigma^2$ in (3.30) and (3.31), respectively. Therefore, if there are

[5]The solution that minimizes the free energy in (3.29) is equivalent to the thermodynamically dominant solution in statistical physics [10, 11].

[6]Convergence to another fixed-point solution is possible if LAMA is initialized sufficiently close to such a fixed point [114].

multiple fixed points to (3.30), then LAMA is, in general, sub-optimal and does not necessarily converge to the fixed-point solution with minimal free-energy.

Before we delve into the optimality analysis of LAMA, we note that the fixed-point analysis for LAMA with noise mismatch is more involved as it requires finding fixed points for the coupled fixed-point equations in (3.31). Hence, we focus on the case $N_0 = N_0^{\text{post}}$.

### 3.4.3 When Does LAMA Achieve the Same Performance as IO Data Detector?

We note that the performance of LAMA (in the large-system limit) is fully described by the SE framework. However, characterizing the performance of the IO data detector is a non-trivial task. Although an analysis via the replica method [11] was recently proved to be correct under mild assumptions [115], a verification of the assumptions still requires extensive work for each prior distribution. Therefore, to establish optimality of LAMA, we first introduce an additional assumption to characterize the performance of the IO data detector, and then show that under this assumption, LAMA achieves the same data detection performance as the IO data detector.

We define a specific example of a large-sparse limit that will be used for our analysis of LAMA. The general definition of large-sparse limit is provided in [113].

**Definition 5.** *We define the large-sparse limit as the following procedure: First, start by defining a binary-valued matrix* $\mathbf{B} \in \{0,1\}^{B \times U}$. *Pick a constant* $\Gamma \leq U$ *and generate*

*each entry $B_{b,u}$ as an i.i.d. Bernoulli random variable with probability $\Gamma/U$. Define a normalization constant $\Gamma_u = \sum_{b=1}^{B} B_{b,u}$ for each $u = 1, \ldots, U$. Then, generate the channel matrix $\mathbf{H}$ with each entry being i.i.d. $H_{b,u} \in \mathcal{CN}(0, 1/\Gamma_u)$ if $B_{b,u} = 1$ and 0 otherwise. Based on this construction of $\mathbf{H}$ for a fixed $\Gamma$, we define the large-sparse limit when we first let $B, U \to \infty$ with $U/B = \beta$. Then, we let $\Gamma \to \infty$.*

We note that the large-system limit corresponds to the case when we first set $\Gamma = U$ and then let $B, U \to \infty$ with $U/B = \beta$. However, we will assume that we first fix a constant $\Gamma < U$, and then let $B, U \to \infty$; this formulation of the large-sparse limit is needed to prevent the factor graph for the input-output relation in (2.1) from having short cycles [113]. We need an additional assumption to establish optimality of LAMA. We assume that exchanging the order of the large-system limit still holds true for cSE:

**(A3)** We assume that cSE for LAMA remains valid in the large-sparse limit.

With Definition 5 and (A3), we will now establish optimality of LAMA in two parts. First, we show that in the large-sparse limit, BP achieves the same performance as the IO data detector and the input-output relation is asymptotically decoupled into AWGN channels with equal decoupled noise variance. Second, we show that LAMA achieves the same noise variance as that given by BP using state evolution. Since the input-output relation is decoupled into AWGN channels and LAMA achieves the lowest (unique) decoupled noise variance, LAMA achieves the same detection performance as the IO data detector. We show the first part by [113, Thm. 4]:

**Theorem 7.** *Assume the large-sparse limit and the system ratio $\beta_{\mathrm{BP}}$ is chosen such that the fixed-point solution of BP $\eta_{\mathrm{BP}}$ to (3.28) is unique. Then, BP achieves the same*

*performance as the IO data detector. In addition, the posterior distribution of each user after BP converges to that given by an AWGN channel with variance $N_0/\eta_{\mathrm{BP}}$.*

Theorem 7 shows that in the large-sparse limit and for unique fixed points, one can use BP to achieve the same performance as IO data detector. The proof in [113, Sec. V] uses a sandwiching argument between genie-aided BP and classical BP to achieve IO performance. Interestingly, the posterior distribution of each transmit stream after BP converges to that given by an AWGN channel. In addition, the noise variance of the equivalent AWGN channel can be characterized by solving a fixed-point equation (3.28); this fixed-point equation coincides exactly to that given by the replica method shown in [11]. Now that we have shown that BP achieves IO performance and characterized the decoupling of AWGN, we now establish optimality of LAMA.

**Corollary 8.** *Assume the large-system limit and $\beta_{\mathrm{LAMA}} = \beta_{\mathrm{BP}}$ from Theorem 7. Then, LAMA decouples the MIMO system into parallel AWGN channels with variance $\sigma^2_{\mathrm{IO}}$, which is a unique fixed-point solution to (3.30) with $\sigma^2_{\mathrm{IO}} = N_0/\eta_{\mathrm{BP}}$ from Theorem 7.*

The proof of Corollary 8 follows from first noting that (3.28) and (3.30) are equal. Hence, since $\beta_{\mathrm{LAMA}} = \beta_{\mathrm{BP}}$, LAMA has a unique fixed-point solution to (3.30) given by $\sigma^2_{\mathrm{IO}}$ which is equivalent to $N_0/\eta_{\mathrm{BP}}$. Since LAMA decouples the MIMO system into parallel AWGN channels [82] and the decoupled variances are equal, LAMA achieves the same performance as the IO data detector. In Section 3.4.5, we provide conditions for which there is exactly one (unique) fixed point with minimum effective noise variance $\sigma^2$.

### 3.4.4 Optimality Conditions for LAMA with No Noise : Exact Recovery Thresholds (ERTs)

We start by analyzing LAMA in a noiseless setting and for $N_0 = N_0^{\text{post}} = 0$. We provide sharp bounds on the system ratio $\beta = U/B$, which guarantee exact recovery of an unknown transmit signal $\mathbf{s}_0$ in the large-system limit. We show that if $\beta < \beta_{\mathcal{O}}^{\max}$, where $\beta_{\mathcal{O}}^{\max}$ is the so-called *exact recovery threshold (ERT)*, then LAMA perfectly recovers $\mathbf{s}_0$. Note that the ERT depends on the constellation $\mathcal{O}$ and resembles to the phase-transition behavior observed in sparse signal recovery [98, 116, 117]; the key difference is that LAMA operates with dense vectors.

We will show in Theorem 10 that if $\beta < \beta_{\mathcal{O}}^{\max}$, there exists a unique fixed point at $\sigma^2 = 0$ to the fixed-point equation in (3.30). The unique fixed point at $\sigma^2 = 0$ implies that the effective noise variance output for the decoupled AWGN channel will be zero. Therefore, the output from the non-linear MMSE estimate from LAMA will be $\mathsf{F}(\mathbf{s}_0, \sigma^2) = \mathbf{s}_0$ from (3.19), and hence LAMA perfectly recovers $\mathbf{s}_0$. For $\beta \geq \beta_{\mathcal{O}}^{\max}$, perfect recovery cannot be guaranteed.[7] To make this behavior explicit, we need the following technical result with proof in Appendix A.1.8.

**Lemma 9.** *Fix the constellation set $\mathcal{O}$ and let $\mathbb{V}\text{ar}_S[S]$ be finite. Then, there exists a non-negative gap $\sigma^2 - \Psi(\sigma^2, \sigma^2) \geq 0$ with equality if and only if $\sigma^2 = 0$. As $\sigma^2 \to 0$, we have MSE $\Psi(\sigma^2, \sigma^2) \to 0$; as $\sigma^2 \to \infty$, we have the the MSE $\Psi(\sigma^2, \sigma^2) \to \mathbb{V}\text{ar}_S[S]$.*

For a finite value of $\mathbb{V}\text{ar}_S[S]$, Lemma 9 shows that we have $\Psi(\sigma^2, \sigma^2) < \sigma^2$ for all $\sigma^2 > 0$. Now, suppose that for some $\beta > 1$, $\beta\Psi(\sigma^2, \sigma^2) < \sigma^2$ also holds for all

[7] We assume the initialization as given in Algorithm 2. LAMA may recover the original signal for $\beta \geq \beta_{\mathcal{O}}^{\max}$ if initialized sufficiently close to the optimal fixed point; see [114] for a discussion.

$\sigma^2 > 0$. Then, as long as $\beta > 1$ is not too large to also ensure $\beta\Psi(\sigma^2, \sigma^2) < \sigma^2$, for all $\sigma^2 > 0$, there will only be a *single* fixed point at $\sigma^2 = 0$. Therefore, LAMA is able to perfectly recover the original signal $\mathbf{s}_0$ by Theorem 2 since the unique fixed point at $\sigma^2 = 0$ implies that $\Psi(\sigma^2, \sigma^2) = 0$. Leveraging the gap between $\Psi(\sigma^2, \sigma^2)$ and $\sigma^2$ will allow us to find the exact recovery threshold (ERT) of LAMA for values of $\beta > 1$. For the fixed (discrete) constellation $\mathcal{O}$, the largest value of $\beta$ that ensures $\beta\Psi(\sigma^2, \sigma^2) < \sigma^2$ is precisely the ERT.

**Definition 6.** *Fix $\mathcal{O}$ and let $N_0 = N_0^{\text{post}} = 0$. Then, the exact recovery threshold (ERT) that enables perfect recovery by LAMA is defined by*

$$\beta_{\mathcal{O}}^{\max} = \min_{\sigma^2 \geq 0} \left\{ \left( \frac{\Psi(\sigma^2, \sigma^2)}{\sigma^2} \right)^{-1} \right\}. \tag{3.32}$$

We are now ready to establish perfect recovery with $\beta_{\mathcal{O}}^{\max}$; the proof is given in Appendix A.1.9.

**Theorem 10.** *Let $N_0 = N_0^{\text{post}} = 0$ and $\mathbf{H}$ satisfy (A2). Fix the constellation $\mathcal{O}$. If $\beta < \beta_{\mathcal{O}}^{\max}$, then LAMA perfectly recovers $\mathbf{s}_0$ in (2.1) in the large-system limit.*

We emphasize that for a given constellation $\mathcal{O}$, the ERT $\beta_{\mathcal{O}}^{\max}$ can be computed numerically from (3.32), where $\Psi(\sigma^2, \sigma^2)$ is given by Theorem 2. We emphasize that the signal variance, $\mathbb{V}\text{ar}_S[S]$ does not have an impact on the ERT as the MSE function $\Psi(\sigma^2, \sigma^2)$ and $\sigma^2$ both scale linearly with $\mathbb{V}\text{ar}_S[S]$. In Section 3.4.5, we extend our analysis to the noisy case.

Table 3.1: Summary of (Sub-)Optimality Regimes of LAMA

| | $\beta < \beta_{\mathcal{O}}^{\min}$ | $\beta_{\mathcal{O}}^{\min} \leq \beta \leq \beta_{\mathcal{O}}^{\max}$ | $\beta_{\mathcal{O}}^{\max} < \beta$ |
|---|---|---|---|
| $N_0 < N_0^{\min}(\beta)$ | *optimal* | *optimal* | suboptimal |
| $N_0^{\min}(\beta) \leq N_0 \leq N_0^{\max}(\beta)$ | *optimal* | (sub-)optimal[a] | suboptimal |
| $N_0^{\max}(\beta) < N_0$ | *optimal* | *optimal* | *optimal* |

[a]For some constellations, there may exist intervals in $[N_0^{\min}(\beta), N_0^{\max}(\beta)]$ where LAMA is still optimal; an example is shown in Fig. 3.6.

### 3.4.5 Optimality Conditions for LAMA With Noise: Maximum Ratio Thresholds (MRTs)

We develop optimality conditions of LAMA in the presence of noise, and we focus on mismatch-free case as the associated optimality conditions allow for an elegant analysis.[8]

In the presence of noise ($N_0 > 0$), exact recovery is no longer guaranteed. Nevertheless, if LAMA converges to a unique fixed-point, then we obtain the same error-rate performance as the IO data detector. In such situations, we call LAMA to be optimal. Furthermore, if multiple fixed-points exist, we call the fixed-point with minimum effective noise variance the *optimal fixed point*, whereas all other fixed points are called *suboptimal fixed points*.

In essence, there exist three different regimes for LAMA (see Table 3.1), which depend on the system ratio $\beta$:

(i) $\beta < \beta_{\mathcal{O}}^{\min}$ : if $\beta$ is smaller than the so-called *minimum recovery threshold* (MRT) $\beta_{\mathcal{O}}^{\min}$, then LAMA is *always* guaranteed to converge to the unique fixed point

[8]The mismatch-free case requires us to identify all fixed points of (3.30), whereas mismatch case requires the identification of all fixed points to the *coupled* fixed-point equations in (3.31). A detailed analysis of optimality conditions for LAMA with noise variance mismatch is left for future work.

(with minimal $\sigma^2$), i.e., the LAMA delivers IO data detection performance irrespective of the noise variance $N_0$.

(ii) $\beta_{\mathcal{O}}^{\min} \leq \beta \leq \beta_{\mathcal{O}}^{\max}$ : if $\beta$ is larger or equal to than the MRT, but smaller than or equal to the ERT, then multiple fixed points exist. In this case, optimality of LAMA depends on the noise variance $N_0$. If the noise variance $N_0$ is *larger* than the so-called *maximum guaranteed noise variance* $N_0^{\max}(\beta)$, then LAMA converges to the unique fixed point. Similarly, if the noise variance $N_0$ is strictly *smaller* than the so-called *minimum critical noise* $N_0^{\min}(\beta)$, then LAMA converges to the optimal fixed point. However, if $N_0 \in [N_0^{\min}(\beta), N_0^{\max}(\beta)]$, then LAMA converges, in general, to a sub-optimal fixed point.[9] We also note that for some constellations, there may exist intervals in $[N_0^{\min}(\beta), N_0^{\max}(\beta)]$ in which LAMA remains to be optimal. This behavior is shown in Fig. 3.6. Furthermore, as $\beta \to \beta_{\mathcal{O}}^{\max}$, the minimum critical noise $N_0^{\min}(\beta) \to 0$, which implies that LAMA is optimal when $N_0 > N_0^{\max}(\beta)$.

(iii) $\beta > \beta_{\mathcal{O}}^{\max}$ : If $\beta$ exceeds the ERT, then LAMA is optimal if $N_0 > N_0^{\max}(\beta)$. For all other values of $N_0$, LAMA converges, in general, to a sub-optimal fixed point.

In order to make these three regimes more explicit, we require the following definition.

**Definition 7.** *Fix the constellation $\mathcal{O}$ and let $N_0^{\text{post}} = N_0$. Then, the minimum recovery threshold (MRT) $\beta_{\mathcal{O}}^{\min}$ is defined as follows:*

$$\beta_{\mathcal{O}}^{\min} = \min_{\sigma^2 \geq 0} \left\{ \left( \frac{\mathrm{d}\Psi(\sigma^2, \sigma^2)}{\mathrm{d}\sigma^2} \right)^{-1} \right\}. \tag{3.33}$$

[9] We note that LAMA can still be optimal if it was initialized close to the optimal fixed point [1], but we exclude this case from our analysis.

By the definition of the MRT, it is easy to observe that the fixed point of (3.30) is unique for all system ratios $\beta < \beta_{\mathcal{O}}^{\min}$, as $\beta \frac{\mathrm{d}\Psi(\sigma^2,\sigma^2)}{\mathrm{d}\sigma^2} < 1$ for all values of $\sigma^2$. The following lemma establishes an intuitive relationship between MRT and ERT; the proof is given in Appendix A.1.10.

**Lemma 11.** *The MRT never exceeds the ERT.*

Lemma 11 shows that if the system ratio $\beta$ is less than MRT, i.e., $\beta < \beta_{\mathcal{O}}^{\min}$, then LAMA is not only optimal but also perfect recovery is possible in noiseless settings. We next define the minimum critical and maximum guaranteed noise variance, $N_0^{\min}(\beta)$ and $N_0^{\max}(\beta)$, that determine boundaries for the optimality regimes when $\beta \geq \beta_{\mathcal{O}}^{\min}$.

**Definition 8.** *Fix the system ratio $\beta \in [\beta_{\mathcal{O}}^{\min}, \beta_{\mathcal{O}}^{\max}]$. Then, the* minimum critical *noise variance $N_0^{\min}(\beta)$ that ensures convergence to the optimal fixed-point is defined by*

$$N_0^{\min}(\beta) = \min_{\sigma^2 \geq 0} \left\{ \sigma^2 - \beta \Psi\left(\sigma^2, \sigma^2\right) : \beta \frac{\mathrm{d}\Psi(\sigma^2,\sigma^2)}{\mathrm{d}\sigma^2} = 1 \right\}.$$

**Definition 9.** *Fix the system ratio $\beta \geq \beta_{\mathcal{O}}^{\min}$. Then, the* maximum guaranteed *noise variance $N_0^{\max}(\beta)$ that ensures convergence to the optimal fixed-point is defined by*

$$N_0^{\max}(\beta) = \max_{\sigma^2 \geq 0} \left\{ \sigma^2 - \beta \Psi\left(\sigma^2, \sigma^2\right) : \beta \frac{\mathrm{d}\Psi(\sigma^2,\sigma^2)}{\mathrm{d}\sigma^2} = 1 \right\}.$$

Note that as $\beta \to \beta_{\mathcal{O}}^{\max}$, the minimum critical noise decreases to $N_0^{\min}(\beta) \to 0$. To see this, consider the case when $\beta = \beta_{\mathcal{O}}^{\max}$, so that there exists a $\sigma_\star^2 > 0$ such that $\beta_{\mathcal{O}}^{\max} \Psi(\sigma_\star^2, \sigma_\star^2) = \sigma_\star^2$. It is clear that $\beta_{\mathcal{O}}^{\max} \frac{\mathrm{d}\Psi(\sigma^2,\sigma^2)}{\mathrm{d}\sigma^2}\Big|_{\sigma^2=\sigma_\star^2} = 1$ and hence $N_0^{\min}(\beta_{\mathcal{O}}^{\max}) = \sigma_\star^2 - \beta_{\mathcal{O}}^{\max} \Psi(\sigma_\star^2, \sigma_\star^2) = 0$.

Before we proceed with the analysis for optimality regimes of LAMA, we present Lemma 12 (with proof in Appendix A.1.11) that shows how the fixed-point $\sigma^2$ decreases with $N_0$ as $N_0 \to 0$.

**Lemma 12.** *Fix the constellation $\mathcal{O}$ and let $\beta < \beta_{\mathcal{O}}^{\max}$. Denote $\sigma^2$ as the largest fixed-point solution of LAMA with noise variance $N_0$. Then, as $N_0 \to 0$, we have $\sigma^2 \to 0$. In addition, we have $\lim_{N_0 \to 0} \frac{\sigma^2}{N_0} = 1$.*

Lemma 12 shows that not only the fixed-point solution $\sigma^2$ of LAMA goes to 0 as $N_0 \to 0$, but also decreases linearly as $\lim_{N_0 \to 0} \frac{\sigma^2}{N_0} = 1$. We now proceed to the optimality regime analysis. We recall that all the zero-crossing points of the function

$$g(\sigma^2, \beta, N_0, \mathcal{O}) = N_0 + \beta \Psi(\sigma^2, \sigma^2) - \sigma^2 \tag{3.34}$$

correspond to all the fixed points of the cSE of LAMA. We will frequently refer to the function in (3.34) for our optimality analysis of LAMA.

Figures 3.3–3.5 illustrates our optimality analysis for a large MIMO system with QPSK normalized to $E_s = 1$. We plot the function (3.34) depending on the effective noise variance $\sigma^2$ and for different system ratios $\beta$. The cases $\beta < \beta_{\mathcal{O}}^{\min}$, $\beta \in [\beta_{\mathcal{O}}^{\min}, \beta_{\mathcal{O}}^{\max}]$, and $\beta > \beta_{\mathcal{O}}^{\max}$ are shown in Fig. 3.3, Fig. 3.4, and Fig. 3.5, respectively. The special case of $\beta = 1$ in the noiseless setting $N_0 = 0$ for (3.34) corresponds to the solid blue line, along with the corresponding (unique) fixed point at the origin. In the following three paragraphs, we discuss the three operation regimes of LAMA.

**(i)** $\beta < \beta_{\mathcal{O}}^{\min}$

In this region, the cSE of LAMA always converges to the unique, optimal fixed point. For $\beta < \beta_{\mathcal{O}}^{\min}$, the slope of (3.34) is strictly-negative. Hence, as (3.34) is always decreasing, there exists exactly one unique fixed point for the cSE of LAMA

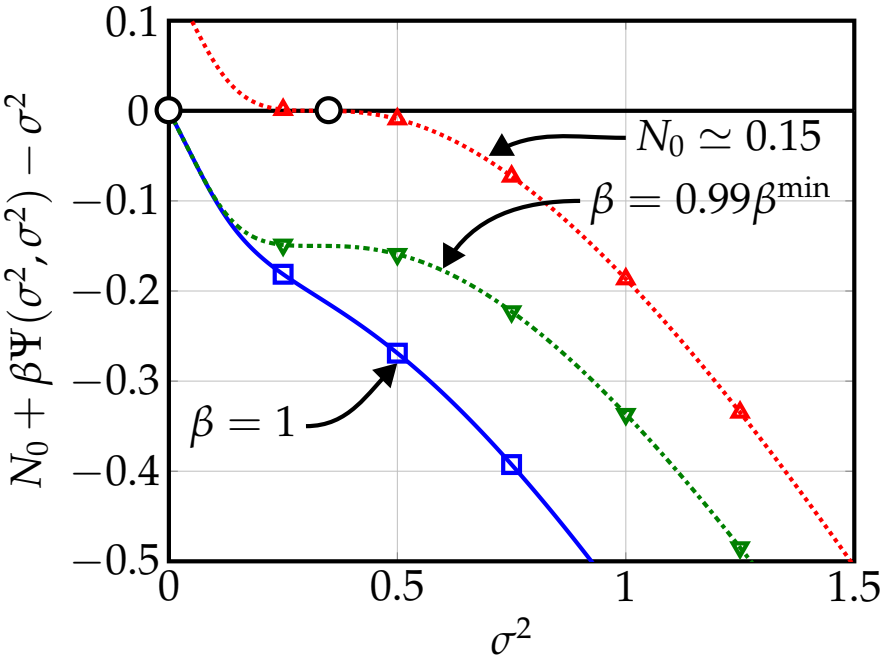



Figure 3.3: Function (3.34) for $\beta < \beta^{\min}_{\text{QPSK}}$: LAMA always converges to the unique, optimal fixed point, irrespective of the noise variance $N_0$.

regardless of the noise variance $N_0$. Thus, LAMA achieves IO performance. The green dash-dotted and red dotted line in Fig. 3.3 show (3.34) for $\beta < \beta^{\min}_{\text{QPSK}}$ with $N_0 = 0$ and $N_0 \simeq 0.15$, respectively. In both cases, we see that the cSE of LAMA converges to the unique fixed point.

**(ii)** $\beta^{\min}_{\mathcal{O}} \leq \beta \leq \beta^{\max}_{\mathcal{O}}$

In this region, the cSE of LAMA converges to the unique, optimal fixed point if $N_0 < N_0^{\min}(\beta)$ or if $N_0 > N_0^{\max}(\beta)$ and consequently, LAMA achieves IO performance in both of these regimes. The green dash-dotted line, cyan dashed line, and magenta dotted line in Fig. 3.4 show (3.34) for $\beta^\star = (\beta^{\min}_{\text{QPSK}} + \beta^{\max}_{\text{QPSK}})/2$ with $N_0 = 0$, $N_0 > N_0^{\max}(\beta^\star)$ and $N_0 < N_0^{\min}(\beta^\star)$, respectively. We note that for the three cases the fixed point is unique, labeled in Fig. 3.4 by a circle. The red, dotted line in Fig. 3.4 shows (3.34) with $\beta^\star$ for noise $N_0 \in [N_0^{\min}(\beta^\star), N_0^{\max}(\beta^\star)]$. In this

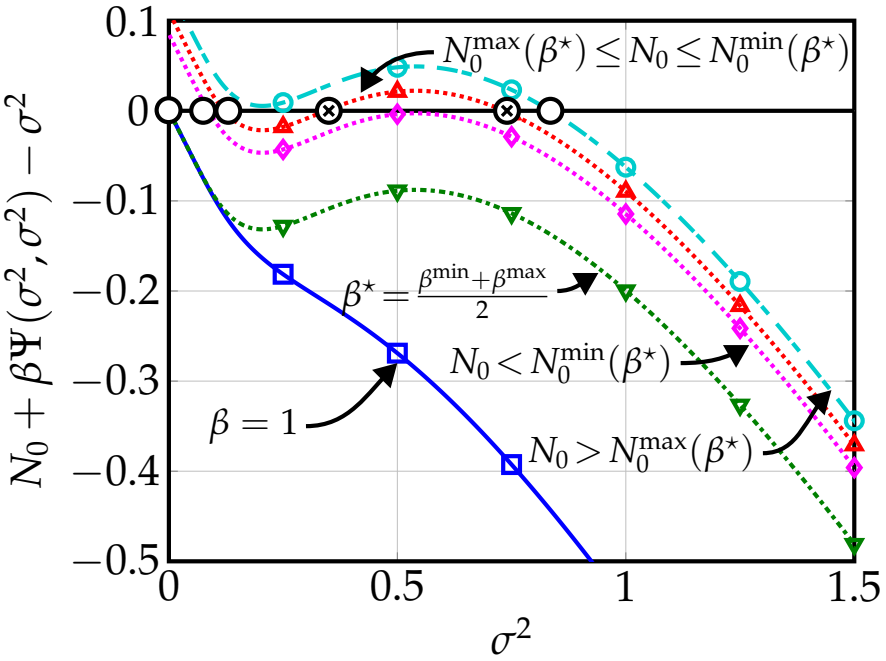



Figure 3.4: Function (3.34) for $\beta \in [\beta^{\min}_{\text{QPSK}}, \beta^{\max}_{\text{QPSK}}]$: For $\beta \in [\beta^{\min}_{\text{QPSK}}, \beta^{\max}_{\text{QPSK}}]$, we have two regimes for which LAMA converges to an optimal fixed point: (i) $N_0 < N_0^{\min}(\beta)$ and (ii) $N_0 > N_0^{\max}(\beta)$. The situation $\beta^\star = \frac{\beta^{\min}_{\text{QPSK}}+\beta^{\max}_{\text{QPSK}}}{2}$ with (i) $N_0 < N_0^{\min}(\beta^\star)$ is shown with a purple dotted curve and (ii) $N_0 > N_0^{\max}(\beta^\star)$ is shown shown with a cyan dashed curve; we see that LAMA exhibits a single (and hence, optimal) fixed point. However, if $N_0 \in [N_0^{\min}(\beta), N_0^{\max}(\beta)]$, which is shown with a red dotted curve, the cSE of LAMA exhibits multiple fixed points and hence, LAMA is no longer IO.

case, however, we observe that the cSE of LAMA converges to the rightmost suboptimal fixed point labeled by the crossed circle $\otimes$. Hence, LAMA is able to achieve IO performance if $N_0^{\min}(\beta) \le N_0 \le N_0^{\max}(\beta)$.

**(iii)** $\beta > \beta^{\max}_{\mathcal{O}}$

In this region, the cSE of LAMA converges to the unique, optimal fixed point when $N_0 > N_0^{\max}(\beta)$ and consequently, achieves IO performance. Unlike the previous case for $\beta^{\min}_{\mathcal{O}} \le \beta \le \beta^{\max}_{\mathcal{O}}$, for which LAMA has *two* regions of optimality, $N_0 > N_0^{\max}(\beta)$ and $N_0 < N_0^{\min}(\beta)$, for $\beta > \beta^{\max}_{\mathcal{O}}$, LAMA has only one optimal

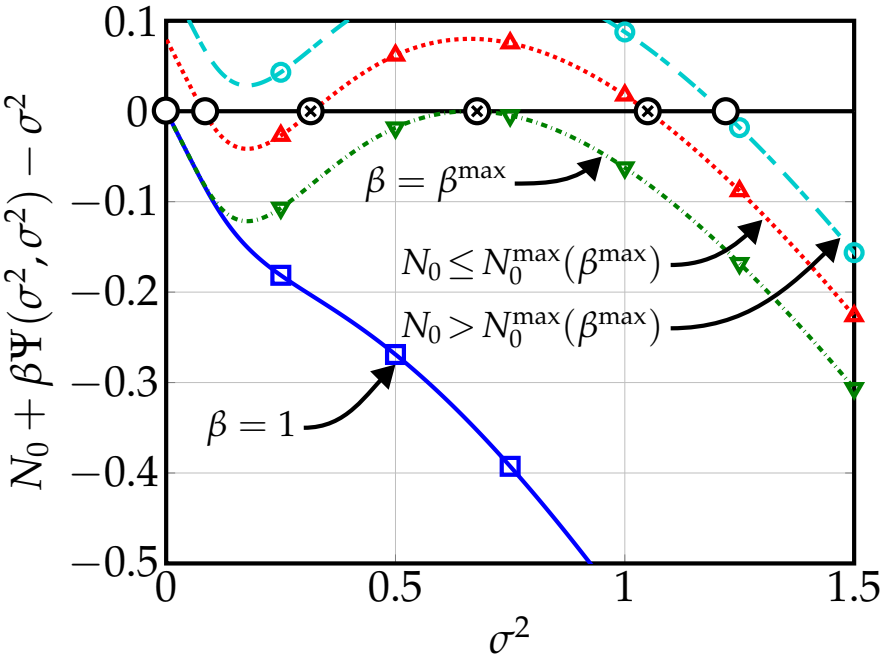



Figure 3.5: Function (3.34) for $\beta > \beta^{\max}_{\text{QPSK}}$: For $\beta > \beta^{\max}_{\text{QPSK}}$, the cSE of LAMA converges to a suboptimal fixed point in the noiseless case $N_0 = 0$, which is shown in green. However, when $N_0 > N_0^{\max}(\beta)$, the cSE of LAMA converges to the optimal fixed point, which can be seen in the cyan dashed curve. If $N_0 \leq N_0^{\min}(\beta)$, then the cSE of LAMA, shown in red dotted curve, has multiple fixed points and thus, is no longer IO.

region: $N_0 > N_0^{\max}(\beta)$. As $\beta \to \beta^{\max}_{\mathcal{O}}$, the low noise $N_0 < N_0^{\min}(\beta)$ (or high SNR regime) region of optimality disappears because $N_0^{\min}(\beta) \to 0$ as $\beta \to \beta^{\max}_{\mathcal{O}}$ from (3.32). The green, dash-dotted line and red, dotted lines in Fig. 3.5 show (3.34) for $\beta = \beta^{\max}_{\text{QPSK}}$ with $N_0 = 0$ and $0 < N_0 \leq N_0^{\max}(\beta)$, respectively. We observe that the cSE of LAMA converges to the suboptimal fixed point when $\beta = \beta^{\max}_{\text{QPSK}}$ even with $N_0 = 0$. The cyan, dashed line refers to $\beta = \beta^{\max}_{\text{QPSK}}$ with $N_0 > N_0^{\max}(\beta)$. While the noiseless case enables the cSE of LAMA to converge to the suboptimal fixed point, we observe that for high noise (or equivalently low SNR), the cSE of LAMA is able to achieve IO performance. Therefore, if $\beta > \beta^{\max}_{\mathcal{O}}$, then LAMA achieves IO performance whenever the noise variance exceeds the maximum guaranteed noise variance $N_0^{\max}(\beta)$.

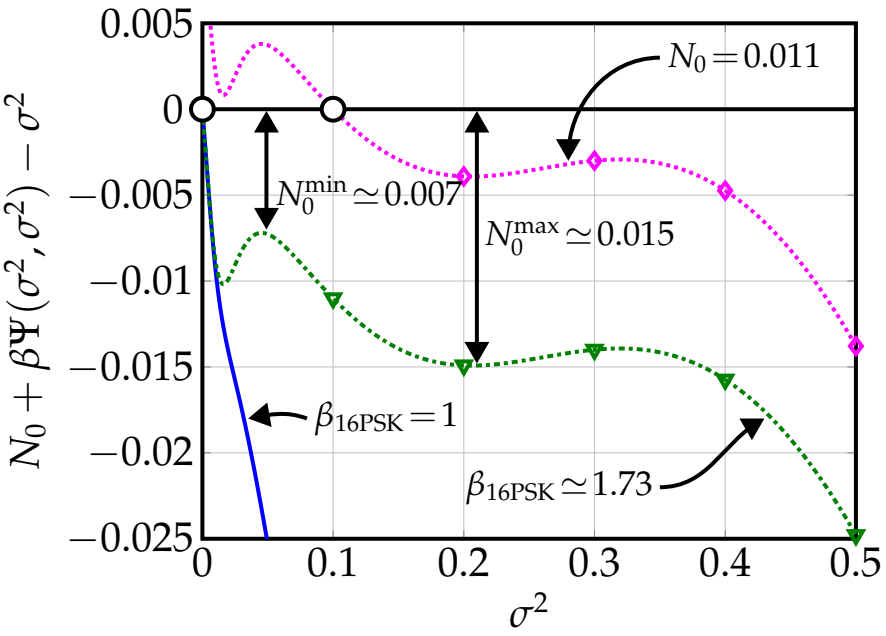



Figure 3.6: Function (3.34) for 16-PSK and $\beta = 1.73 \in [\beta^{\min}_{16\text{-PSK}}, \beta^{\max}_{16\text{-PSK}}]$: For 16-PSK and $\beta = 1.73 \in [\beta^{\min}_{16\text{-PSK}}, \beta^{\max}_{16\text{-PSK}}]$, $N_0^{\min}(\beta)$ and $N_0^{\max}(\beta)$ is computed to be 0.007 and 0.015 respectively. For 16-PSK, there exists regions where $N_0 \in [N_0^{\min}(\beta), N_0^{\max}(\beta)]$ and LAMA still achieves IO performance.

We also note that for some constellations, the cSE of LAMA may recover the optimal fixed point for $\beta \in [\beta^{\min}_{\mathcal{O}}, \beta^{\max}_{\mathcal{O}}]$ in some noise variance intervals $[N_0^{\min}(\beta), N_0^{\max}(\beta)]$. An example case for $\beta = 1.73$ with 16-PSK is shown in Fig. 3.6, where cSE of LAMA recovers the unique fixed-point with $N_0 = 1.1 \cdot 10^{-2} \in [N_0^{\min}(\beta), N_0^{\max}(\beta)]$. These intervals exist for some constellations because in addition to $\sigma^2$ that result $N_0^{\min}(\beta)$ and $N_0^{\max}(\beta)$, there are multiple values of $\sigma^2$ that satisfy $\frac{\mathrm{d}}{\mathrm{d}\sigma^2} g(\sigma^2, \beta, N_0, \mathcal{O}) = 0$, where $g(\sigma^2, \beta, N_0, \mathcal{O})$ is defined in (3.34). As a result, there exist intervals between $N_0^{\min}(\beta)$ and $N_0^{\max}(\beta)$ that the cSE of LAMA has one (optimal) fixed point. In such regions, LAMA enables IO performance. We finally note that the MRT $\beta^{\min}_{\mathcal{O}}$ and ERT $\beta^{\max}_{\mathcal{O}}$ do not depend on the signal variance $\mathbb{V}\mathrm{ar}_S[S]$. In contrast, the critical noise levels $N_0^{\min}(\beta)$ and $N_0^{\max}(\beta)$ depend on $\mathbb{V}\mathrm{ar}_S[S]$.

### 3.4.6 Decomposing Complex-Valued Systems

We now analyze whether the cSE of LAMA with complex-valued constellations can equivalently be characterized by a real-valued SE with a real-valued constellation. We note that while the loading factor limits were given in [118] and [10] respectively, these results were pertinent to BPSK with real-valued systems and no results were given for other constellations.

We note that the standard way of dealing with complex-valued systems is via the real-valued decomposition (see footnote 1). This approach, however, violates the independent assumption on the MIMO channel. Since LAMA operates directly on the complex plane, no transformation into the real-valued domain is required. Nevertheless, we now provide conditions for which the complex-valued problem can be exactly characterized by a corresponding real-valued problem. For our analysis, we require the following definition.

**Definition 10.** *For all $s \in \mathcal{O}$, express $s$ as $s = a + ib$, where $a \in \mathrm{Re}(\mathcal{O})$, $b \in \mathrm{Im}(\mathcal{O})$. Then, the constellation $\mathcal{O}$ is called* separable *if $p(s) = p(a)p(b)$ holds for all $s \in \mathcal{O}$ and $\mathrm{Re}(\mathcal{O}) = \mathrm{Im}(\mathcal{O})$.*

For example, $M^2$-QAM with equally likely symbols is separable. In contrast, $M^2$-PSK is not separable (except for QPSK) as the real and imaginary parts dependent. We now present a result that allows us to transform the complex-valued cSE equations in (3.11) and (3.12) into equivalent real-valued SE equations; the proof is given in Appendix A.1.12.

We do so by first deriving the MSE and variance function for the real-valued separable constellation on a real-valued system. Then, we will link the complex-valued and real-valued SE by noting that cSE is *exactly* characterized by the

real-valued SE.

**Theorem 13.** *Let the constellation $\mathcal{O}$ be separable. Define $S_{\mathrm{R}} = \mathrm{Re}(S)$ and denote the real-part of $\mathcal{O}$ as $\mathcal{O}^{\mathrm{R}}$. Define $\mathsf{F}^{\mathrm{R}}$ and $\mathsf{G}^{\mathrm{R}}$ as the message mean and variance function, respectively, with $\mathrm{Re}(S) \sim p(\mathrm{Re}(S))$. Then, the complex-valued posterior mean and variance functions can be computed separately in real and imaginary parts as:*

$$\begin{aligned}\mathsf{F}(S, \tau^2) &= \mathsf{F}^{\mathrm{R}}(\mathrm{Re}(S), \tau^2/2) + \mathrm{i}\mathsf{F}^{\mathrm{R}}(\mathrm{Im}(S), \tau^2/2) \\ \mathsf{G}(S, \tau^2) &= \mathsf{G}^{\mathrm{R}}(\mathrm{Re}(S), \tau^2/2) + \mathsf{G}^{\mathrm{R}}(\mathrm{Im}(S), \tau^2/2).\end{aligned}$$

*Also define the MSE function $\Psi$ and the variance function $\Phi$ for the real-valued prior $\mathrm{Re}(S)$ on a real-valued system as:*

$$\begin{aligned}\Psi^{\mathrm{R}}(\sigma^2, \gamma^2) &= \mathbb{E}_{S_{\mathrm{R}}, N_{\mathbb{R}}}\left[\left(\mathsf{F}^{\mathrm{R}}(S_{\mathrm{R}} + \sigma N_{\mathbb{R}}, \gamma^2) - S_{\mathrm{R}}\right)^2\right], \\ \Phi^{\mathrm{R}}(\sigma^2, \gamma^2) &= \mathbb{E}_{S_{\mathrm{R}}, N_{\mathbb{R}}}\left[\mathsf{G}^{\mathrm{R}}(S_{\mathrm{R}} + \sigma N_{\mathbb{R}}, \gamma^2)\right].\end{aligned}$$

*Then we have the following relation for $\Psi$ and $\Phi$ between the complex-valued constellation $\mathcal{O}$ and the real-valued constellation $\mathcal{O}^{\mathrm{R}}$:*

$$\Psi(\sigma^2, \gamma^2) = 2\Psi^{\mathrm{R}}\left(\frac{\sigma^2}{2}, \frac{\gamma^2}{2}\right), \quad \Phi(\sigma^2, \gamma^2) = 2\Phi^{\mathrm{R}}\left(\frac{\sigma^2}{2}, \frac{\gamma^2}{2}\right).$$

*Therefore, the cSE recursions in (3.11) and (3.12) are given by:*

$$\sigma_t^2 = N_0 + \beta\Psi(\sigma_t^2, \gamma_t^2) = N_0 + 2\beta\Psi^{\mathrm{R}}\left(\frac{\sigma_t^2}{2}, \frac{\gamma_t^2}{2}\right), \tag{3.35}$$

$$\gamma_t^2 = N_0^{\mathrm{post}} + \beta\Phi(\sigma_t^2, \gamma_t^2) = N_0^{\mathrm{post}} + 2\beta\Phi^{\mathrm{R}}\left(\frac{\sigma_t^2}{2}, \frac{\gamma_t^2}{2}\right). \tag{3.36}$$

We note that LAMA operates *simultaneously* on complex-valued signals by reducing $\sigma_t^2$ each iteration in both real and imaginary parts independently; this can be seen by noting that since $\mathcal{O}$ is separable, $\Psi^{\mathrm{R}}$ is identical for both the real and imaginary parts of $\mathcal{O}$. In addition, Theorem 13 shows that if $\mathcal{O}$ is separable,

Table 3.2: ERTs $\beta_{\mathcal{O}}^{\max}$, MRTs $\beta_{\mathcal{O}}^{\min}$ and the critical noise levels $N_0^{\min}(\beta_{\mathcal{O}}^{\min})$ and $N_0^{\max}(\beta_{\mathcal{O}}^{\max})$ for LAMA with common PSK, PAM, and QAM constellations

| Constellation | | $\beta_{\mathcal{O}}^{\min}$ | $N_0^{\min}(\beta_{\mathcal{O}}^{\min})$ | $\beta_{\mathcal{O}}^{\max}$ | $N_0^{\max}(\beta_{\mathcal{O}}^{\max})$ |
|---|---|---|---|---|---|
| $\mathbb{C}$ system | $\mathbb{R}$ system | | | | |
| BPSK | – | 2.951 | $3.00 \cdot 10^{-1}$ | 4.171 | $2.43 \cdot 10^{-1}$ |
| QPSK | BPSK | 1.475 | $1.50 \cdot 10^{-1}$ | 2.086 | $1.22 \cdot 10^{-1}$ |
| 16-QAM | 4-PAM | 0.983 | $3.00 \cdot 10^{-2}$ | 1.363 | $2.45 \cdot 10^{-2}$ |
| 64-QAM | 8-PAM | 0.842 | $7.14 \cdot 10^{-3}$ | 1.157 | $5.87 \cdot 10^{-3}$ |
| 256-QAM | 16-PAM | 0.786 | $1.77 \cdot 10^{-3}$ | 1.075 | $1.45 \cdot 10^{-3}$ |
| 8-PSK | – | 1.458 | $4.44 \cdot 10^{-2}$ | 1.804 | $3.83 \cdot 10^{-2}$ |
| 16-PSK | – | 1.473 | $1.14 \cdot 10^{-2}$ | 1.801 | $9.95 \cdot 10^{-3}$ |
| 64-PSK | – | 1.474 | $7.23 \cdot 10^{-4}$ | 1.801 | $8.39 \cdot 10^{-3}$ |
| 256-PSK | – | 1.474 | $4.52 \cdot 10^{-5}$ | 1.801 | $8.39 \cdot 10^{-3}$ |

then the cSE can be transformed into a real-valued SE, hence validating the relation between the complex-valued constellation and the equivalent real-valued representation. This transformation implies that for certain constellations, the message mean $\mathsf{F}$ and variance function $\mathsf{G}$ can be computed (often more efficiently) in parallel for real and imaginary dimensions.

We note that in [11] Guo and Verdú used a real-valued decomposition and the replica method for analyzing the performance of complex-valued signals for separable constellations and concluded that the error performance for complex signals is exactly same as that of real-valued system with transmit energy halved. Theorem 13 supports this conclusion. Moreover, we emphasize that the cSE holds for *general* constellations, such as higher-order PSK constellations, and LAMA can be used for data detection in such cases.

### 3.4.7 ERT, MRT, and Critical Noise Levels

The ERT, MRT, as well as the critical noise levels $N_0^{\min}(\beta)$ and $N_0^{\max}(\beta)$ for common constellations and for real-valued as well as complex-valued systems are summarized in Table 3.2. We assume equally likely priors with the constellation sets normalized to $\mathbb{V}\text{ar}_S[S] = E_s = 1$. We note that the calculations of ERT and MRT for the simplest case with BPSK involve computations of logistic-normal integrals for which no closed-form expressions are known [119] but approximations exist [119–121]. The results in Table 3.2 were obtained via numerical integration to compute the MSE function $\Psi(\sigma^2, \sigma^2)$.[10] Next Lemma shows that for real- and separable complex-valued constellations, the ERT and MRT are identical for real- and complex-valued systems, respectively; a short proof is given in Appendix A.1.13. For an example, BPSK for real-valued systems and QPSK for complex-valued systems have identical ERT and MRT of 1.475 and 2.086, respectively.

**Lemma 14.** *Fix a separable constellation $\mathcal{O}$ and denote $\beta_{\mathbb{C}}^{\min}$, $\beta_{\mathbb{C}}^{\max}$ and $\beta_{\mathbb{R}}^{\min}$, $\beta_{\mathbb{R}}^{\max}$ as MRT and ERT of the complex and real-valued constellation, respectively. Also, denote the critical noise levels $N_{0,\mathbb{C}}^{max}(\beta)$, $N_{0,\mathbb{C}}^{min}(\beta)$, $N_{0,\mathbb{R}}^{max}(\beta)$, and $N_{0,\mathbb{R}}^{max}(\beta)$ for the complex and real-valued constellation, respectively. Then, $\beta_{\mathbb{C}}^{\min} = \beta_{\mathbb{R}}^{\min}$, $\beta_{\mathbb{C}}^{\max} = \beta_{\mathbb{R}}^{\max}$, $N_{0,\mathbb{C}}^{max}(\beta) = 2N_{0,\mathbb{R}}^{max}(\beta)$, and $N_{0,\mathbb{C}}^{min}(\beta) = 2N_{0,\mathbb{R}}^{min}(\beta)$.*

Lemma 14 implies that optimality results for $M^2$-QAM in a complex system with equally likely transmit symbols (shown in Table 3.2) are the same for a real-valued $M$-PAM system. Moreover, between BPSK and QPSK in a complex system, we observe that all the thresholds differ by a factor of 2, which is expected. As shown in the second row of Table 3.2 for QPSK with complex noise, or a

[10] We used MATLAB's `integral` and `integral2` commands with $\texttt{AbsTol} = \texttt{RelTol} = 10^{-12}$.

real-valued BSPK system (with real noise) the ERT is $\beta_{\text{QPSK}}^{\max} \approx 2.0855$, which corresponds exactly to the maximum loading factor for the IO data detector established in [10,11]. Moreover, the MRT for QPSK is given as $\beta_{\text{QPSK}}^{\min} \approx 1.4752$ [10].[11] The critical noise values in Table 3.2 refer to complex constellations as the critical noise values can be easily computed for the real constellation by Lemma 14.

The MRTs for 16-QAM and 64-QAM indicate that small system ratios $\beta < 1$ are necessary to guarantee that LAMA achieves IO performance. For instance, we require $\beta \leq \beta_{\text{64-QAM}}^{\min} \approx 0.8424$, i.e. $U \leq 0.8424B$, to ensure that LAMA solves (2.3) for 64-QAM. As $\beta \to \beta_{\text{64-QAM}}^{\max} \approx 1.1573$, LAMA is only optimal in settings in which the noise level is rather high, i.e., where $N_0 > N_0^{\max}(\beta_{\text{64-QAM}}^{\max}) \approx 5.868 \cdot 10^{-3}$, or, equivalently, when $\mathsf{SNR} < 22.9495$ dB. From Table 3.2, we see that higher-order QAM or PSK constellations can be decoded optimally by LAMA in massive MIMO as one typically assumes $B \gg U$. We also observe that as $M$ increases for $M$-PSK, $\beta_{\mathcal{O}}^{\min}$ and $\beta_{\mathcal{O}}^{\max}$ approaches to 1.4741 and 1.8005, respectively.

## 3.5 Numerical Asymptotic Results

We now provide numerical results for LAMA, discuss practical implementation aspects, and highlight the pros and cons. In what follows, we use the average received SNR defined in (3.1).

[11] Note that $\beta_{\text{QPSK}}^{\min}$ Tanaka provided in [10] is 1.49, whereas we obtain a slightly more accurate value 1.4752.

### 3.5.1 Achievable Rates and Error-Rate Performance

As detailed in Section 3.3.3, the output of LAMA enables one to represent each transmit stream by a single-input single-output AWGN channel with a equal noise variance $\sigma_t^2$ that can be computed via the cSE Theorem 2. Therefore, the performance of LAMA in the large-system limit can be characterized by analyzing a single AWGN channel.

Figures 3.7 and 3.8 show the achievable rate and symbol error rate for LAMA after 100 iterations for various system ratios $\beta$. While an infinite number of iterations would guarantee LAMA to converge to a fixed point solution, our results show that much fewer than 100 iterations are required for LAMA to converge; we will further discuss this aspect in Section 3.5.2.

Fig. 3.7 shows the achievable rate of the decoupled AWGN channel per transmit stream for LAMA, for various system ratios $\beta$. For small values of $\beta$, e.g. $\beta = 0.1$, the achievable rate of LAMA approaches to that of an AWGN channel, which agrees with Lemma 6. We observe that the performance gap between LAMA and that of an AWGN channel increases with $\beta$. In particular, when $\beta = \beta_{\text{QPSK}}^{\text{min}}$, we see a sudden transition in the achievable rate of LAMA to the achievable rate of an AWGN channel, which occurs approximately at 10 dB, which also happens for $\beta = \beta_{\text{QPSK}}^{\text{mid}} = \frac{1}{2}(\beta_{\text{QPSK}}^{\text{max}} + \beta_{\text{QPSK}}^{\text{max}})$. This transition occurs exactly at the SNR regime for which the noise variance $N_0$ becomes smaller than $N_0^{\text{min}}(\beta)$, which was shown to ensure convergence of LAMA to the unique optimal fixed point (cf. Section 3.4.5). For $\beta \to \beta_{\text{QPSK}}^{\text{max}}$, we see that the achievable rate does not converge to that of an interference-free AWGN channel, irrespective of the SNR regime; this agrees with the perfect recoverability result in the large-system limit shown in Theorem 10 for ERT $\beta_{\text{QPSK}}^{\text{max}}$.

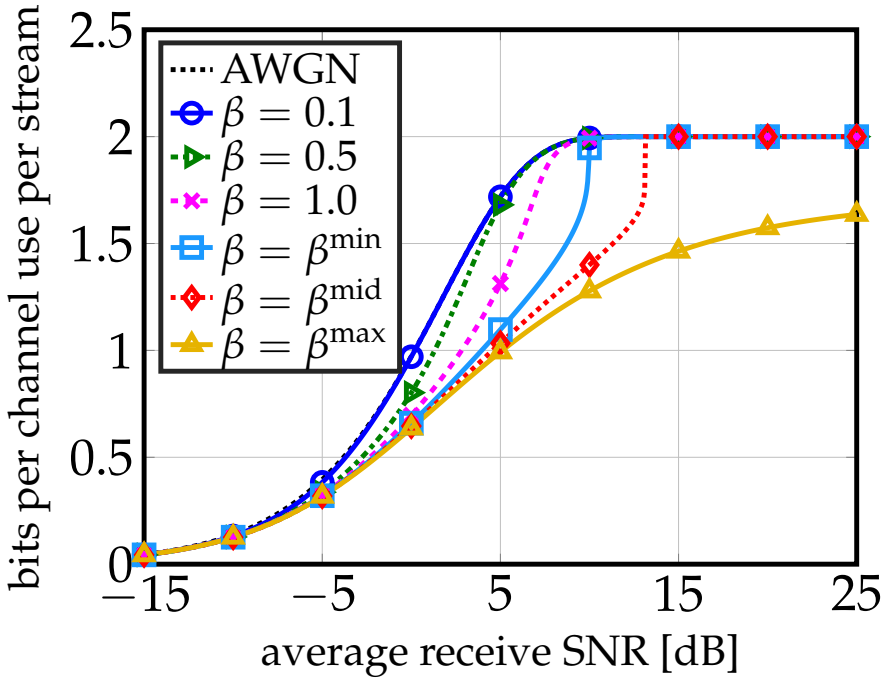



Figure 3.7: Achievable rate (in bits per channel use per stream) of LAMA for different system ratios $\beta$ with QPSK constellation. The sharp transitions occur at $\beta > \beta^{\min}_{\text{QPSK}}$ when the $\mathsf{SNR} = \beta \frac{E_s}{N_0}$ with $N_0$ equaling the critical noise variance $N_0^{\min}(\beta)$.

Fig. 3.8 shows the symbol error rate (SER) of LAMA. Similar to the achievable rate in Fig. 3.7, the SER for $\beta = 0.1$ for LAMA in a MIMO system approaches that of an interference-free AWGN channel. For $\beta > \beta^{\min}$, we observe a waterfall behavior where the SER quickly drops and approaches that of an interference-free AWGN channel; this happens at exactly the point where the noise variance is smaller than the minimum critical noise $N_0^{\min}(\beta^{\min}_{\text{QPSK}})$. We note that this waterfall behavior is consistent with the SNR regime that caused an upwards jump in the achievable rate curve shown in Fig. 3.7. When $\beta = \beta^{\max}_{\text{QPSK}}$, we observe an SER floor at about 0.08; this is due to the fact that as $\mathsf{SNR} \to \infty$, the cSE of LAMA always converges to a suboptimal fixed point shown in Fig. 3.5.

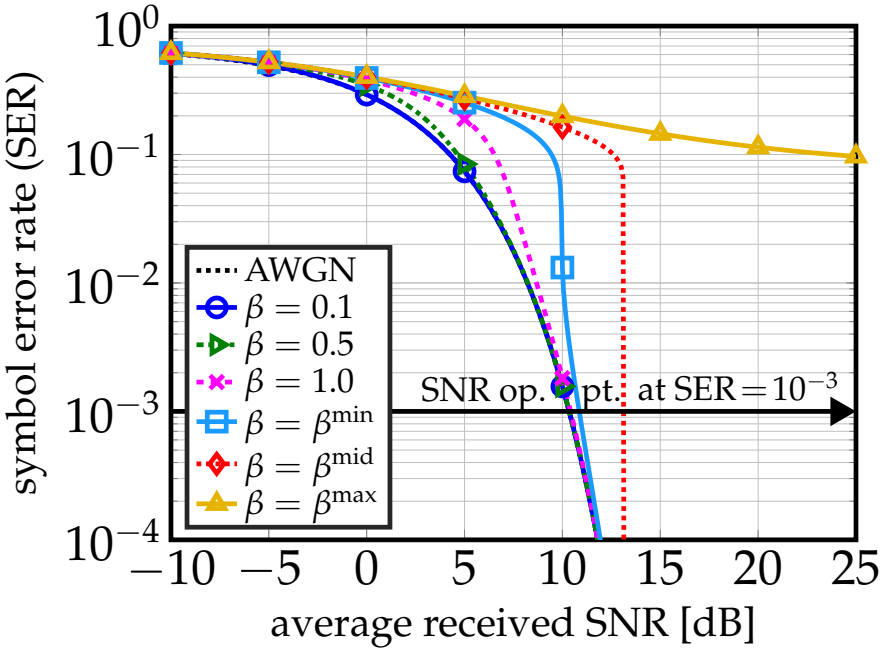



Figure 3.8: Symbol error rate (SER) of LAMA for different system ratios $\beta$ with QPSK constellation. The sharp transitions occur at $\beta > \beta^{\min}_{\text{QPSK}}$ when the $\mathsf{SNR} = \beta \frac{E_s}{N_0}$ with $N_0$ equaling the critical noise variance $N_0^{\min}(\beta)$.

### 3.5.2 Performance and Complexity Trade-offs

While only an infinite number of LAMA iterations guarantee the cSE of LAMA in Theorem 2 to converge to a fixed-point of (3.30) and (3.31), one can terminate the algorithm early with the goal of reducing its complexity. A straightforward approach is to terminate Algorithm 2, if the parameter $\hat{\tau}_t^2$ does not improve from one iteration to the next, e.g., if $\hat{\tau}_t^2 \leq \hat{\tau}_{t+1}^2$ is met. Another approach is to terminate LAMA after a predefined number of $I$ iterations. The latter approach not only enables a deterministic throughput (which is critical in hardware implementations), but also enables us to study a fundamental performance/complexity trade-off of LAMA.

Since the cSE analysis is only valid in the large system limit, common complexity measures, such as the number of additions and/or multiplications are

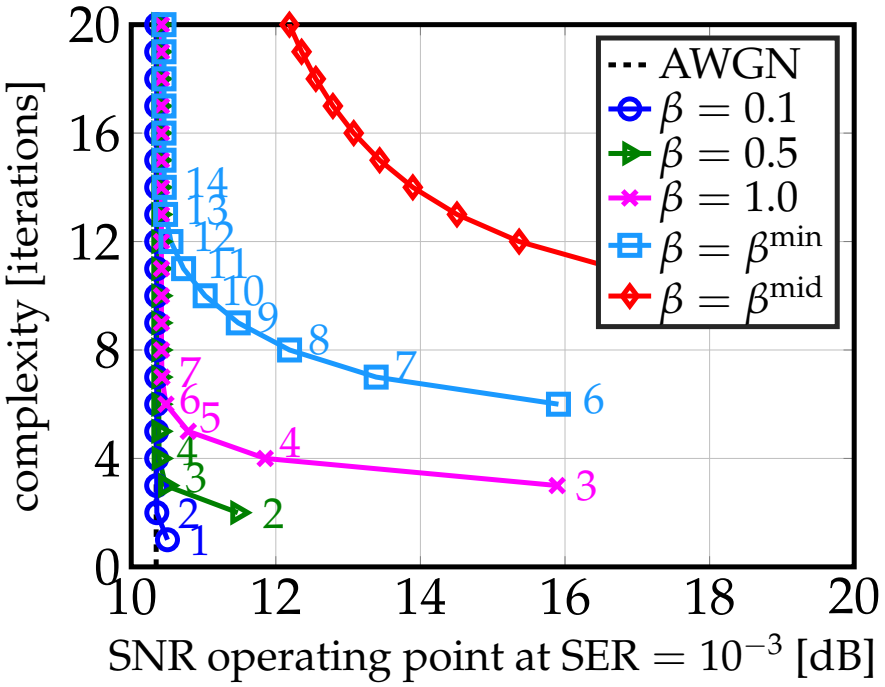



Figure 3.9: Performance/complexity trade-offs of LAMA for different system ratios $\beta$ with QPSK constellation. The dashed lines refer to the SNR operating point for an AWGN channel at $SER = 10^{-3}$ for each $\beta$. The atomic complexity of LAMA required to approach AWGN SNR increases with the system ratio $U/B = \beta$.

not meaningful. Nevertheless, for a given system, we see from Algorithm 2 that the computational workload of LAMA per iteration remains constant. Hence, counting the maximum number of algorithm iterations provides a sensible way of measuring the complexity of LAMA.[12]

**Definition 11.** *The* atomic complexity *of LAMA is defined by the maximum number of algorithm iterations $I$.*

We now study the performance of LAMA depending on the atomic complexity $I$. Put simply, we investigate by how much one can approach the performance of LAMA with infinitely many iterations. We do so by first computing the output

[12] In practice, one can multiply the atomic complexity with the number of arithmetic operations require per iteration; this enables one to obtain an accurate complexity measure that depends on the system configuration.

variance $\sigma_I^2$ of the equivalent AWGN channel for a fixed complexity $I$, and then computing the associated SER.

We first discuss the convergence speed of LAMA to its fixed-point solution. The following result, with proof in Appendix A.1.14, reveals that if $\beta < \beta_{\mathcal{O}}^{\min}$, then LAMA not only has a unique fixed point solution, but also converges exponentially fast; this ensures that LAMA achieves near-IO performance with a small number of iterations.

**Lemma 15.** *Assume the initialization of LAMA as in Algorithm 2. If $\beta < \beta_{\mathcal{O}}^{\min}$, then regardless of the noise variance $N_0$, LAMA converges exponentially fast to its unique fixed-point solution $\sigma_\star^2$.*

Fig. 3.9 shows the required SNR to achieve SER of $10^{-3}$ for every iteration of LAMA for various systems ratios $\beta$ in the large-system limit. The colored dashed lines correspond to the SNR required to achieve an SER of $10^{-3}$ in an interference-free AWGN channel, which we call "AWGN SNR." For $\beta = 0.1$, $\beta = 0.5$, only three and five iterations are required for LAMA to closely approach the AWGN SNR. We observe that as $\beta$ decreases, the number of iterations required to reach SNR of SER $10^{-3}$ also decreases. This observation is in accordance with Lemma 6, where we demonstrated that in the extreme case where $\beta \to 0$, one iteration (matched filter detection) is sufficient to converge to the AWGN SNR. As $\beta$ increases, we start to see the performance differences between LAMA and that of an interference-free AWGN channel. For $\beta = \beta_{\text{QPSK}}^{\min}$, the SNR operating point of LAMA closely approaches the AWGN SNR after 15 iterations at a small performance loss (about 0.1 dB), which is visible from the SER plot in Fig. 3.8. The differences between the SNR operating point of LAMA and AWGN SNR are more pronounced when $\beta = (\beta_{\text{QPSK}}^{\min} + \beta_{\text{QPSK}}^{\max})/2$, as the SNR operating point

of LAMA converges to 13.5 dB after about 90 iterations, which is 0.6 dB higher than the AWGN SNR of 12.9 dB. For $\beta = \beta_{\text{QPSK}}^{\max}$, the complexity of LAMA is not shown as it floors to an SER of approximately 0.08 and hence, never achieves the target SER of $10^{-3}$.

### 3.5.3 Performance in Finite-Dimensional Systems

Since the design of LAMA heavily relies on the large system limit, there are no optimality guarantees for finite-dimensional settings. For conventional, small-scale MIMO systems (with 8 antennas or less), the large-system assumption leads to a significant performance loss because (i) the statistics of $\hat{\mathbf{z}}^t = \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t$ are not Gaussian and hence, (ii) the correct statistics of the Gaussian term $\hat{\mathbf{z}}^t$ cannot be tracked in the LAMA algorithm. The problem that arises in finite-dimensional systems becomes evident if we keep $\beta = 1$ and increase $\mathsf{SNR} \to \infty$ for a small system. We see that LAMA exhibits in an SER floor (see Fig. 3.13 for a $128 \times 128$ 16-QAM system). We note that this SER floor lowers as the system's dimension increases. The performance loss of AMP-based algorithms for small-sized systems has been investigated in [122, 123].

In order to mitigate LAMA's performance loss in finite dimensional systems, one can use estimators as opposed to the original message variance function in (3.6) to estimate $\sigma_t^2$ each iteration. For estimators in LAMA to work universally when the antenna configurations are *both* small and big, we need estimators of $\sigma_{t+1}^2$, which we will denote as $\hat{\sigma}_{t+1}^2$, that not only lower the error floor at high SNR in small antenna systems, but also converges to the true effective noise variance $\sigma_{t+1}^2$ in large antenna systems.

In [86], a series of estimators have been proposed for AMP in the context of sparse recovery. We adopt the same approach for LAMA for the case $N_0^{\text{post}} = N_0$, where instead of computing the average of the exact message variance function as (3.10), we estimate the variance of the Gaussian estimate $\hat{\mathbf{z}}^t = \hat{\mathbf{s}}^t + \mathbf{H}^{\text{H}}\mathbf{r}^t$ by:

$$\hat{\sigma}_{t+1}^2 = \frac{1}{B}\left\|\mathbf{r}^t\right\|_2^2. \tag{3.37}$$

Fig. 3.13 shows the performance of (3.37) for LAMA in an $128 \times 128$ system with 16-QAM. We observe a decrease in the SER floor in high SNR regime compared to the original LAMA without the estimator with no performance loss in the low SNR regime.

### 3.5.4 Extension to General Channel Matrices H

It is important to note that one of the limiting assumptions underlying AMP (and hence, for LAMA) is that the entries of the channel matrix $\mathbf{H}$ are i.i.d. zero-mean Gaussian or complex Gaussian with variance $1/B$ for AMP and complex-valued AMP respectively. In practical systems, however, the BS antennas may exhibit correlation and uneven power profiles, especially in multi-user scenarios, which makes LAMA less robust in these scenarios. To address these limitations, Rangan [59] has developed Generalized AMP (GAMP), which extends AMP to arbitrary input and output noise distributions for real-valued systems, and can operate in channels with different power profiles. We note that in the large system limit with $\mathbf{H}$ distributed according to (A2) with Gaussian noise, GAMP and AMP are equivalent. In addition, a modified GAMP that uses damping technique was proposed in [124] to cope with non-zero mean, low-rank channels. The damping technique slows certain algorithmic parameter updates, but does so at the cost

of increased iterations of the algorithm. Vila and Schniter furthermore included expectation-maximization into GAMP in [91,100], which further improves the performance of AMP-based methods in finite-dimensional systems. Recently, reference [90] introduced vector AMP, which further generalizes GAMP to arbitrary matrices. Generalized AMP has been used for practical MIMO-OFDM systems [103] with variations introduced in [91,100,102] to increase the detection performance for a finite-dimensional system. Reference [103] primarily focused on simulations, whereas this chapter concentrates on theoretical analysis in the large-system limit via the state-evolution framework.

### 3.5.5 Simulation Results

We show simulation results for LAMA for a QPSK constellation in Figure 3.10 and Fig. 3.11. We plot the finite-dimensional and asymptotic performance of LAMA from numerical simulations and SE equations, respectively. Figure 3.10 demonstrates that LAMA (i) is able to achieve the error-rate performance of the individually optimal detector for a square MIMO system (i.e., $B = U$) in the large-system limit, and (ii) closely approaches the error-rate performance of the IO data detector in finite-dimensional systems. Furthermore, we can accurately characterize the performance/complexity trade-offs without the need for expensive system simulations; see Fig. 3.11 for an illustration.

Figures 3.12 and 3.13 show simulation results for large MIMO systems with 16-QAM. We fix the number of BS antennas to 128 and the number of user antennas to 64 and 128. We compare the performance of LAMA to unbiased linear MMSE detection, another message-passing-based receiver, i.e., channel

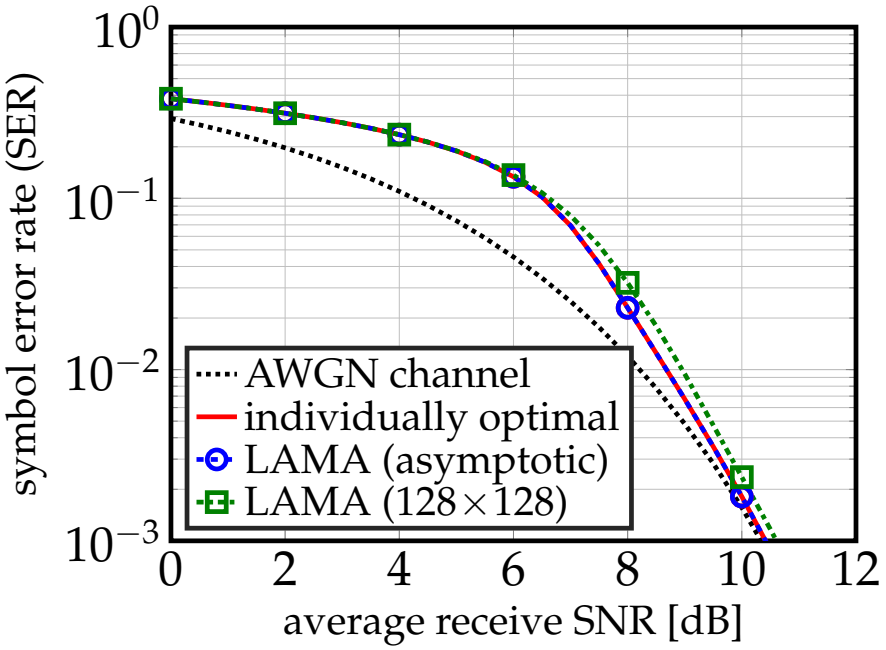



Figure 3.10: Symbol error rate (SER) of LAMA in large MIMO with a square i.i.d. Gaussian system matrix and QPSK modulation in the large-system limit compared to the optimal SER and the SER of an AWGN channel. LAMA achieves the same error-rate performance as the IO data detector and approaches AWGN performance for sufficiently large SNR values; we also see that LAMA closely approaches the theoretical performance limits for finite dimensions (i.e., for a $128 \times 128$ MIMO system).

hardening-exploiting message passing (CHEMP) [111][13], and IO data detection bound obtained by the cSE in the large-system limit.

For the $128 \times 64$ system in Fig. 3.12, LAMA performs very close to the IO bound with only 8 iterations. We note that CHEMP [111] with 8 iterations performs worse than the linear MMSE detection, but approaches the performance of LAMA at 15 iterations. Note that in the large-system limit for a system-ratio of $\beta = 64/128$, $\beta < \beta^{\min}_{16\text{QAM}}$, so LAMA achieves IO data detection performance for any noise variance $N_0$.

[13]We note that CHEMP has no theoretical performance guarantees and was primarily developed for massive MIMO, i.e., $B \gg U$ or small $\beta$.

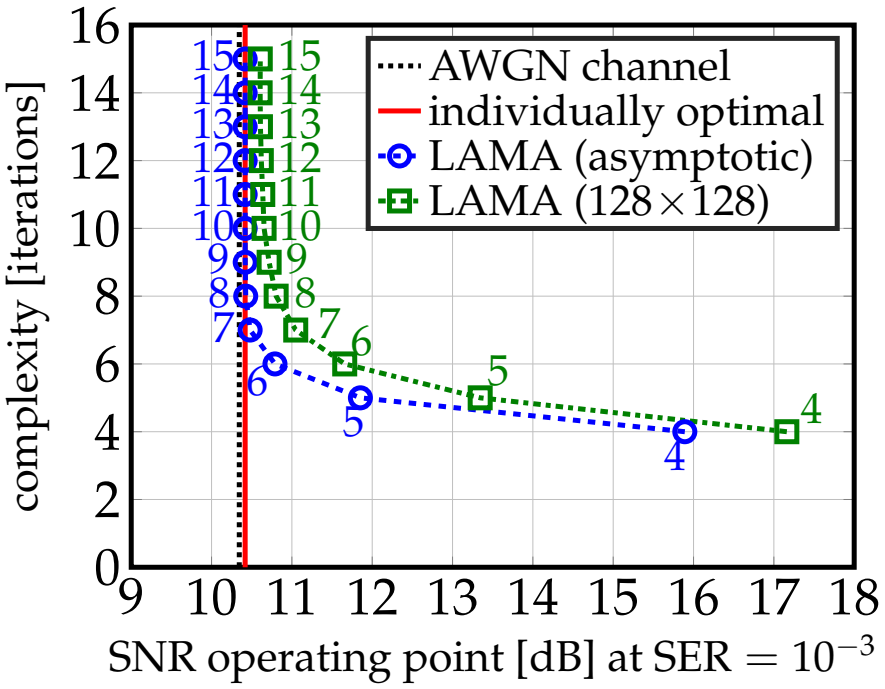



Figure 3.11: Performance/complexity trade-off of LAMA in large MIMO with a square i.i.d. Gaussian system matrix and QPSK modulation in the large-system limit (analytical) and for finite dimensions (simulated); a small number of LAMA iterations is sufficient to approach the theoretical performance limits.

For the $128 \times 128$ system in Fig. 3.13, LAMA with the estimator in (3.37) exhibits a floor at around $10^{-2}$ SER. We note that LAMA with the estimator reduces the error floor while maintaining the performance at low SNR. Because of the flooring behavior of LAMA in finite dimensions, it performs worse than linear MMSE at high SNR (above 35 dB for this case). LAMA with 20 iterations outperforms CHEMP at the same number of iterations; CHEMP floors at an SER of $10^{-1}$ even after 100 iterations. In the $128 \times 128$ setting, we note that $\beta = 1$ is larger than the ERT, $\beta^{\min}_{16\text{QAM}} \approx 0.9830$, from Table 3.2, so LAMA has two regions of optimality (cf. Table 3.1 for the regions) with $N_0^{\min}(\beta) \approx 0.03$, or SNR around 15 dB. Note that this SNR happens where the sharp "waterfall" appears in the IO bound in Fig. 3.13. We stress that for $\beta = 1$ and $U \to \infty$, the SER of LAMA will converge to that of the IO bound by cSE.

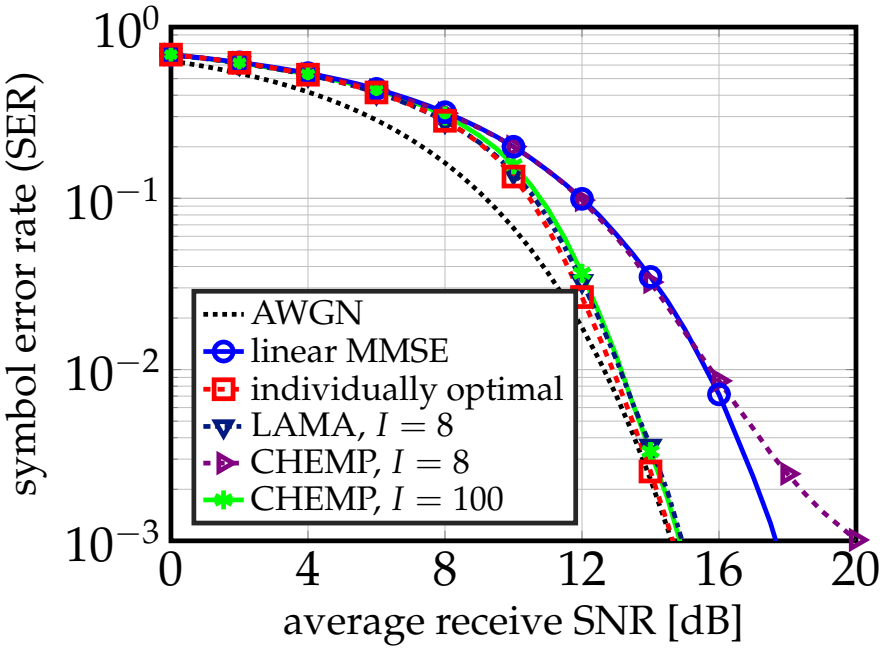



Figure 3.12: Symbol error rate (SER) performance of LAMA for a $128 \times 64$ system with 16-QAM compared to linear MMSE detection and CHEMP [111].

Although we have assumed perfect channel knowledge throughout this chapter to assess the performance of LAMA, we run simulations to see assess the performance of LAMA when there are channel estimation errors. Fig. 3.14 shows the symbol error rate performance of LAMA, CHEMP and linear MMSE. For channel estimation, we used a maximum-likelihood estimate, where we assume orthogonal pilots. We first note that all the data detectors observe a performance degradation of about 3 dB. However, LAMA and CHEMP still outperform linear MMSE. For an example, LAMA and CHEMP achieves SER of $10^{-2}$ at an SNR of 8 dB and 11 dB with perfect and imperfect channel knowledge, whereas linear MMSE needs 13 dB, and 16 dB respectively. Therefore, although all of our analysis pertained to perfect channel knowledge, Fig. 3.14 shows that LAMA still is able to achieve competitive performance when there are channel estimation errors.

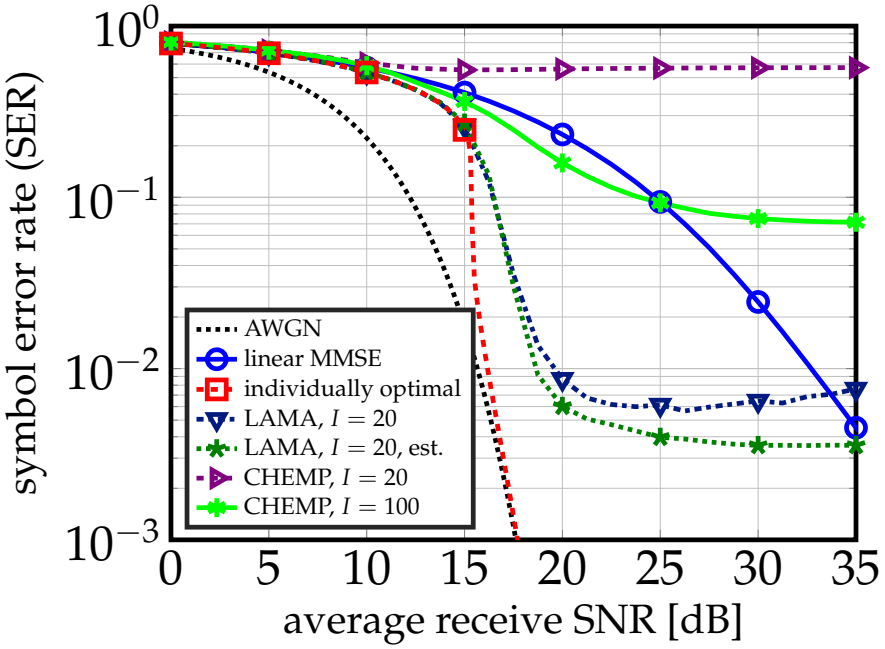



Figure 3.13: Symbol error rate (SER) performance of LAMA with the estimator in (3.37) for a $128 \times 128$ system with 16-QAM to mitigate the performance loss occurring from finite dimension.

## 3.6 LAMA and Prior Art

We now review existing results that are relevant for LAMA and our analysis in Section 3.4.5.

### 3.6.1 BPSK Signaling in Randomly Spread CDMA systems

We show that LAMA for BPSK constellation in randomly spread CDMA systems coincides exactly to the detection algorithm put forth in [71] and the cSE of LAMA without noise variance mismatch is equivalent to that given by IO data detection bound derived from the replica method by Tanaka in [10]. Consider a real-valued randomly-spread CDMA system with equally likely BPSK symbols $\mathcal{O} = \{-1, +1\}$ and the entries of the channel matrix $\mathbf{H}$ are distributed $\mathcal{N}(0, 1/B)$.

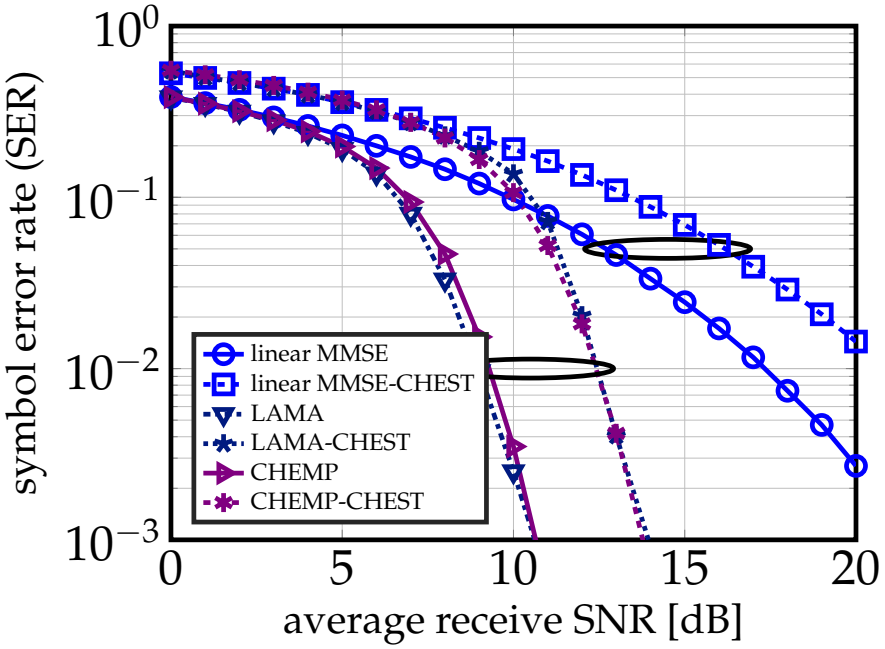



Figure 3.14: Symbol error rate (SER) performance of LAMA in a $12 \times 128$ system with QPSK constellation and channel estimation errors. Even with channel estimation errors, LAMA and CHEMP significantly outperform linear MMSE data detection.

In this case, (3.19) and (3.20) are given by

$$\mathsf{F}(\hat{s}_u, \tau^2) = \tanh\left(\frac{1}{\tau}\hat{s}_u\right), \quad \mathsf{G}(\hat{s}_u, \tau^2) = 1 - \tanh^2\left(\frac{1}{\tau}\hat{s}_u\right).$$

and thus, LAMA corresponds to the following recursion:

$$\begin{aligned}
\hat{\mathbf{s}}^{t+1} &= \tanh\left(\frac{\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{T}}\mathbf{r}^t}{N_0 + \hat{\tau}_t^2}\right) \\
\hat{\tau}_{t+1}^2 &= \beta\left\langle 1 - \tanh^2\left(\frac{\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{T}}\mathbf{r}^t}{N_0 + \hat{\tau}_t^2}\right)\right\rangle \\
\mathbf{r}^{t+1} &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \frac{\hat{\tau}_{t+1}^2}{N_0 + \hat{\tau}_t^2}\mathbf{r}^t,
\end{aligned}$$

with the SE recursion from Theorem 2 given by

$$\sigma_{t+1}^2 = N_0 + \beta\,\mathbb{E}_{S,N_{\mathbb{R}}}\left[\left(\tanh\left(\frac{S + \sigma_t N_{\mathbb{R}}}{\sigma_t^2}\right) - S\right)^2\right], \tag{3.38}$$

where for a fixed $\beta$ and $N_0$, the fixed-point equation is simplified to:

$$\mathbb{E}_{S,N_{\mathbb{R}}}\left[\left(\tanh\left(\frac{S+\sigma_t N_{\mathbb{R}}}{\sigma_t^2}\right)-S\right)^2\right]=1-\int_{\mathbb{R}}\tanh\left(\frac{1}{\sigma_t^2}+\frac{1}{\sigma_t}z\right)\frac{\exp\left(-z^2/2\right)}{\sqrt{2\pi}}\mathrm{d}z,$$

because

$$\begin{aligned}
&\int_{\mathbb{R}}\tanh(\alpha^2+\alpha z)\left(\tanh(\alpha^2+\alpha z)-1\right)\frac{\exp\left(-z^2/2\right)}{\sqrt{2\pi}}\mathrm{d}z\\
&=-2\int_{\mathbb{R}}\frac{1-\exp\left(-2u\right)}{\left(\exp\left(u\right)+\exp\left(-u\right)\right)^2}\exp\left(-\frac{1}{2\alpha^2}(u-\alpha^2)^2\right)\frac{1}{\sqrt{2\pi\alpha^2}}\mathrm{d}u\\
&=-2\exp\left(-\frac{\alpha^2}{2}\right)\int_{\mathbb{R}}\frac{\exp(u)-\exp\left(-u\right)}{\left(\exp\left(u\right)+\exp\left(-u\right)\right)^2}\frac{\exp\left(-\frac{u^2}{2\alpha^2}\right)}{\sqrt{2\pi\alpha^2}}\mathrm{d}u=0.
\end{aligned}$$

Thus, the fixed-point equation is reduced to:

$$\sigma^2=N_0+\beta\int_{\mathbb{R}}\left[1-\tanh\left(\frac{1+\sigma z}{\sigma^2}\right)\right]\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)\mathrm{d}z. \tag{3.39}$$

We note that the fixed point equation in (3.39) coincides exactly to Tanaka's fixed point equation in [10] and the optimal multiuser efficiency in [11] derived using the replica method. Moreover, LAMA coincides exactly to the method developed by Kabashima in 2003 for randomly-spread CDMA with BPSK signaling [71]. Kabashima showed that the algorithm is consistent with the state evolution predictions obtained through numerical simulations. Kabashima's algorithm in [71] was given for BPSK in real-valued systems only; in contrast, LAMA is suitable for general constellations and complex-valued systems, and can be analyzed in the large-system limit.

In Section 3.4.7, we noted that the ERT of a BPSK system for LAMA is computed to be approximately 2.0855, which coincides exactly with Tanaka's recovery threshold in [10], which was computed using the replica method. While we characterized the state of having multiple fixed point solutions by our definitions of MRT and ERT, Tanaka analogized the state of having multiple fixed points

as having coexistence of phases in physical systems. In this context, the MRT $\beta_{\text{BPSK}}^{\min}$ and ERT $\beta_{\text{BPSK}}^{\max}$ corresponds to the boundary in which the instability of the phrases occur, and the boundary where the replica-symmetry solution becomes unstable, breaking the replica-symmetry assumptions [10]. Although an analytical expression of the ERT has been given in [10], an exact characterization of the MRT was not included. Note that our LAMA results generalize Tanaka's results to arbitrary constellations and provide a practical algorithm.

### 3.6.2 Recovery of Antipodal Solutions via Convex Optimization

Recall that from Table 3.2, that a system ratio $\beta$ smaller than 2.0855 is able to perfectly recover a BPSK vector in absence of noise. In this scenario, LAMA is able to determine the unique solution to $\mathbf{y} = \mathbf{H}\mathbf{s}_0$ with $\mathbf{s}_0 \in \{-1, +1\}^U$ if $\mathbf{H}$ is distributed (A2), and $\beta = U/B$ is fixed with $U \to \infty$. A similar scenario was studied in [116, 125], where the authors have provided necessary and sufficient conditions for the recovery of antipodal solutions from $\mathbf{y} = \mathbf{H}\mathbf{s}_0$. In [116], Donoho and Tanner showed that in the large system limit, $\beta < 2$ guarantees the recoverability of the unique signal $\mathbf{s}_0 \in \{-1, +1\}^U$. The same threshold was recovered in [125], by solving the following convex optimization problem [126]:

$$(\mathrm{P}_\infty) \quad \underset{\tilde{\mathbf{s}} \in \mathbb{R}^U}{\text{minimize}} \, \|\tilde{\mathbf{s}}\|_\infty \quad \text{subject to } \mathbf{y} = \mathbf{H}\tilde{\mathbf{s}}.$$

In particular, the solution $\hat{\mathbf{s}}$ to $(\mathrm{P}_\infty)$ corresponds to the antipodal vector $\{-\alpha, +\alpha\}$ for a given $\alpha > 0$ if $\beta < 2$ with high probability [125]. It is interesting to see that $(\mathrm{P}_\infty)$ does not exploit magnitude information (i.e. $\alpha = 1$), whereas LAMA requires this information. Quite surprisingly, the lack of this prior information

only results in a slight improvement in terms of the system ratio $\beta$ that enables perfect recovery from 2 to 2.0855. The error-rate performance of $\hat{\mathbf{s}}$ to $(\mathrm{P}_\infty)$ for was recently investigated in [14,15,127]. We will discuss the connections between the convex-based and LAMA-based methods in Theorem 45.

## 3.7 Conclusions

In this chapter of the thesis, we have developed the complex Bayesian approximate message passing (cB-AMP) framework with a possible mismatch in the postulated noise variance; cB-AMP with appropriate priors enables a derivation of the LAMA data detector. In the large-system limit, we have shown that LAMA decouples large MIMO systems into parallel AWGN channels with identical noise variance across all transmit streams every iteration. Furthermore, cSE has been used to analyze the exact noise variance of the decoupled AWGN channel.

We have derived the specific conditions for which LAMA achieves IO performance. Based on the system ratio $\beta$, there exist three optimality regimes for LAMA, where $\beta \leq \beta_{\mathcal{O}}^{\min}$, $\beta \in (\beta_{\mathcal{O}}^{\min}, \beta_{\mathcal{O}}^{\max})$, and $\beta \geq \beta_{\mathcal{O}}^{\max}$ where the MRT $\beta_{\mathcal{O}}^{\min}$ and ERT $\beta_{\mathcal{O}}^{\max}$ can be computed numerically for LAMA. We have shown both asymptotic and finite-dimensional performance of LAMA through analytical predictions and numerical simulations, which confirm our theoretical results. In addition, we have characterized the convergence behavior of LAMA for system ratios smaller than the ERT. For small system ratios $\beta$, we have shown that LAMA exhibits similar achievable rate and error-rate performance to that of an AWGN channel for a low number of iterations, which makes LAMA an excellent candidate for data detection in massive MIMO systems.

CHAPTER 4

# MISMATCHED DATA DETECTION IN MASSIVE MU-MIMO

## 4.1 Introduction

Data detection in multiple-input multiple-output (MIMO) systems deals with the recovery of the data vector $\mathbf{s}_0 \in \mathcal{O}^U$, where $\mathcal{O}$ is a finite constellation (e.g., QAM or PSK), from the noisy input-output relation $\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}$. In what follows, $U$ and $B$ denotes the number of transmit and receive antennas, respectively, $\mathbf{y} \in \mathbb{C}^B$ is the receive vector, $\mathbf{H} \in \mathbb{C}^{B \times U}$ is the (known) MIMO system matrix, and $\mathbf{n} \in \mathbb{C}^B$ is i.i.d. circularly symmetric complex Gaussian noise. To minimize the symbol error rate, we are interested in solving the following individually-optimal (IO) data detection problem [11, 58, 63]:

$$\text{(IO)} \quad s_\ell^{\text{IO}} = \underset{\tilde{s}_\ell \in \mathcal{O}}{\arg\max}\ p(\tilde{s}_\ell \,|\, \mathbf{y}, \mathbf{H}),\ \ell = 1, \ldots, U,$$

where $s_\ell^{\text{IO}}$ denotes the $\ell$-th IO estimate and $p(\tilde{s}_\ell \,|\, \mathbf{y}, \mathbf{H})$ is the conditional probability density function of $\tilde{s}_\ell^{\text{IO}}$ given the receive vector and the channel matrix.

The IO data detection problem is known to be of combinatorial nature [11, 58, 63] and the use of an exhaustive search or sphere-decoding methods [110] results in prohibitive complexity for systems where $U$ is large. In contrast, our recently proposed algorithm [39], referred to as large MIMO approximate message passing (LAMA), achieves IO performance using a simple iterative procedure in the large-antenna limit, i.e., where we fix the system ratio $\beta = U/B$ and let $U \to \infty$. For practical (finite-dimensional) systems, LAMA was shown to deliver near-IO performance at low computational complexity [39]. Despite these advantages, LAMA requires repeated computations of transcendental functions

that exhibit a high dynamic range. These computations render the design of corresponding high-throughput hardware designs that deploy finite precision (fixed-point) arithmetic a challenging task.

### 4.1.1 Why Should One Use a Mismatched Prior?

In MIMO systems, the true signal prior is generally known, i.e., the transmit constellation is usually known to the receiver. It is therefore natural to ask why the use of a mismatched prior should be useful, especially since the true prior, which leads to the LAMA algorithm Algorithm 2, will minimize the probability of error. To answer this question, we note the following practically-relevant advantages of mismatched detectors: (i) For the LAMA algorithm for equally-likely symbols, the posterior mean function (3.19) corresponds to

$$\mathsf{F}(\hat{s}_u, \hat{\tau}^2) = \frac{\sum\limits_{a \in \mathcal{O}} a \exp\left(-\frac{1}{\hat{\tau}^2}|\hat{s}_u - a|^2\right)}{\sum\limits_{a \in \mathcal{O}} \exp\left(-\frac{1}{\hat{\tau}^2}|\hat{s}_u - a|^2\right)}, \tag{4.1}$$

which exhibits a large dynamic range and requires high arithmetic precision. In fact, even a calculation with double-precision floating point arithmetic becomes numerically unstable for small values of $\hat{\tau}^2$. Hence, the design of high-performance VLSI designs that deploy finite-precision (fixed-point) arithmetic is challenging. Suitably-chosen mismatched priors can alleviate the need for high arithmetic precision and large dynamic range. (ii) While in some situations, the true prior may be unknown to the receivers, some information on the prior may be available (e.g., the energy). We will show in Section 4.3.1 that with mcB-AMP with a mismatched Gaussian prior still enables us to perform the tuning stage in (4.3) with only the knowledge of the energy of the true prior distribution.

### 4.1.2 Relevant Prior Art

Linear data-detection algorithms for MIMO systems, such as zero forcing (ZF) or minimum mean-square error (MMSE), are well-known instances of mismatched data detectors. The performance of such linear detectors in the asymptotic large-system limit has been investigated in [13, 76, 128]. The integral of additional mean-squared estimation error due to the mismatched estimator across all signal-to-noise (SNR) ratios was proven to be the Kullback-Leibler divergence for a single-antenna system [129] and was extended to large MIMO systems in [130]. Another instance of mismatched data detection is the usage of a uniform prior within a hypercube for the recovery of antipodal (e.g., BPSK) signals [15,131–133]. Corresponding theoretical results in [116,125] for noiseless systems revealed that a system ratio of $\beta < 2$ enables perfect signal recovery. The recovery performance for the noisy case was derived recently in [14, 15] by analyzing an $u_\infty$-norm constrained least-squares problem. The theoretical analysis presented in this chapter of the thesis recovers all these results while providing a computationally efficient algorithm.

The mismatched LAMA (LAMA) algorithm proposed in this chapter of the thesis relies upon approximate message passing (AMP) [82, 86, 134], which was developed for sparse signal recovery and compressive sensing. The case of mismatched estimation of sparse signals via AMP was first studied in [135], where the performance of AMP was analyzed when the true prior is unknown. The AMP algorithm in [135, 136] includes a tuning stage for automated parameter selection, which minimizes the output mean-squared error (MSE). The key differences between LAMA as proposed here and the results in [135,136] are that (i) we consider data detection in MIMO systems and (ii) we know the true signal prior

and intentionally select a mismatched prior in order to design hardware-friendly data detection algorithms that enable near-IO performance.

## 4.2 Mismatched Complex Bayesian AMP

We start by presenting a mismatched version of the complex Bayesian approximate message passing (cB-AMP) algorithm (short mcB-AMP), which allows us to use a different prior distribution $\tilde{p}(\tilde{\mathbf{s}})$ that the true signal prior $p(\mathbf{s}_0)$. To minimize the performance loss due to the use of a mismatched prior, we include an optimal tuning stage into the mcB-AMP algorithm.

### 4.2.1 The mcB-AMP Algorithm

Given an i.i.d. prior distribution $p(\mathbf{s}_0) = \prod_{u=1}^{U} p(s_{0,u})$ of the true signal $\mathbf{s}_0$ and a mismatched prior distribution $\tilde{p}(\tilde{\mathbf{s}}) = \prod_{u=1}^{U} \tilde{p}(\tilde{s}_u)$, the mismatched cB-AMP algorithm corresponds to the following iterative procedure:

**Algorithm 3.** *Suppose that* $\mathbf{H}$ *satisfies (A1), then the mismatched cB-AMP is given by the following procedure:*

$$\tilde{\sigma}_t^2 = \frac{1}{B}\|\mathbf{r}^t\|_2^2, \tag{4.2}$$

$$\hat{\tau}_t^2 = \arg\min_{\tau^2} \mathbb{E}_{S_0,N_{\mathbb{C}}}\left[\left|\mathsf{F}^{\mathrm{mm}}(S_0 + \tilde{\sigma}_t N_{\mathbb{C}}, \tau^2) - S_0\right|^2\right], \tag{4.3}$$

$$\hat{\mathbf{s}}^{t+1} = \mathsf{F}^{\mathrm{mm}}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t, \hat{\tau}_t^2), \tag{4.4}$$

$$\mathbf{r}^{t+1} = \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \beta\mathbf{r}^t\left\langle \mathsf{F}'^{\mathrm{mm}}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t, \hat{\tau}_t^2)\right\rangle,$$

*for each iteration* $t = 1, \ldots,$. *The algorithm is initialized by* $\hat{s}_u^1 = \mathbb{E}_{S_0}[S_0]$ *for all*

$u = 1, \ldots, U$, *where* $S_0$ *is a random variable distributed as* $S_0 \sim p(s_0)$, $\mathbf{r}^1 = \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^1$, *and* $\mathsf{F}'^{\mathrm{mm}}$ *is the derivative of* $\mathsf{F}^{\mathrm{mm}}$ *taken by its first argument.*

Similar to cB-AMP shown in Algorithm 1, the functions $\mathsf{F}^{\mathrm{mm}}(s_u, \hat{\tau}^2)$ and $\mathsf{F}'^{\mathrm{mm}}(s_u, \hat{\tau}^2)$ operate element-wise on vectors and the expectation in (4.3) is with respect of the true prior distribution of $S_0$ and $N_{\mathbb{C}} \sim \mathcal{CN}(0, 1)$. The prime difference of mcB-AMP and cB-AMP is that we purposely use a different function for the posterior mean computation for (3.21) that may differ from the true posterior mean function, and the additional tuning steps (4.2) and (4.3). In the following, the function $\mathsf{F}^{\mathrm{mm}}(s_u, \tau^2)$ is the posterior mean with respective to the mismatched prior distribution $\tilde{p}(\tilde{s}_u)$ and variance parameter $\tau^2$ that is defined by:

$$\mathsf{F}^{\mathrm{mm}}(s_u, \tau^2) = \mathbb{E}_{\tilde{S}}[\tilde{S}|s_u] = \int_{\mathbb{C}} \tilde{s} p(\tilde{s}|s_u) \mathrm{d}\tilde{s}, \tag{4.5}$$

where $p(\tilde{s}|s_u)$ is the posterior pdf defined in (3.4), where $p(\tilde{s}|s_u) = \frac{1}{Z} p(s_u|\tilde{s})\tilde{p}(\tilde{s})$ with $p(s_u|\tilde{s}) \sim \mathcal{CN}(\tilde{s}, \tau^2)$ and a normalization constant $Z$. At every iteration, step (4.2) estimates the decoupled noise variance $\sigma_t^2$ (see Section 4.2.2 for a discussion) and step (4.3) tunes the variance parameter $\hat{\tau}_t^2$ based on the estimate for $\sigma_t^2$. The tuning stage in (4.3) ensures that mcB-AMP converges to the solution that minimizes the so-called decoupled noise variance $\sigma_t^2$ for every iteration; see Section 4.2.3 for a proof.

### 4.2.2 Mismatched State Evolution Framework

As discussed in Section 3.2.2 and shown in [1,82], AMP-based algorithms effectively decouple the MIMO system into parallel AWGN channels in the large-system limit, i.e., the quantity $\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t$ can be expressed as $\mathbf{s}_0 + \mathbf{w}^t$, where

$\mathbf{w}^t \sim \mathcal{CN}(0, \sigma_t^2 \mathbf{I}_U)$ and $\sigma_t^2$ is the decoupled noise variance. A key property of AMP-based algorithms is that the decoupled noise variance $\sigma_t^2$ can be tracked *exactly* by the state evolution (SE) framework.

In this section, we detail the mismatched SE framework for the mcB-AMP algorithm shown in Algorithm 3. We note that our mismatched SE framework is a specific instance of the SE framework proved rigorously in [82], with the posterior mean function in (4.5) under the mismatched prior. The mismatched SE for mcB-AMP is shown below in Theorem 16:

**Theorem 16.** *Suppose that* $p(\mathbf{s}_0) = \prod_{u=1}^{U} p(s_{0,u})$ *and mismatched prior* $\tilde{p}(\tilde{\mathbf{s}}) = \prod_{u=1}^{U} \tilde{p}(\tilde{s}_u)$. *Assume the large-system limit, the channel matrix is (A2), and that the posterior mean with respective to the mismatched prior* $\mathsf{F}^{\mathrm{mm}}$ *is a pseudo-Lipschitz function. Then, the decoupled noise variance* $\sigma_{t+1}^2$ *after t iterations of mcB-AMP is given by the following coupled recursion:*

$$\gamma_t^2 = \arg\min_{\gamma^2 \geq 0} \Psi^{\mathrm{mm}}(\sigma_t^2, \gamma^2), \tag{4.6}$$

$$\sigma_{t+1}^2 = N_0 + \beta \Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2), \tag{4.7}$$

*which is initialized by* $\sigma_1^2 = N_0 + \beta \, \mathbb{V}\mathrm{ar}_{S_0}[S_0]$. *Here,* $S_0 \sim p(s_0)$ *and the MSE function is defined by* $\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2) = \mathbb{E}_{S_0, N_{\mathbb{C}}}\left[\left|\mathsf{F}^{\mathrm{mm}}(S_0 + \sigma_t N_{\mathbb{C}}, \gamma_t^2) - S_0\right|^2\right]$, *where the expectation is taken with respective to the true signal* $S_0$ *distributed* $S_0 \sim p(s_0)$ *and* $N_{\mathbb{C}} \sim \mathcal{CN}(0, 1)$.

If the true prior is identical to the mismatched prior, i.e., $p(\mathbf{s}_0) = \tilde{p}(\tilde{\mathbf{s}})$, then cB-AMP in Algorithm 1 selects optimally-tuned parameters which is done in (4.3) (see Section 4.2.3 for detailed discussion). Therefore, mcB-AMP results in the same decoupled noise variance as that given by cB-AMP. The proof of the following result is given in Appendix A.2.1.

**Lemma 17.** *If there is no prior mismatch, i.e., $p(\mathbf{s}_0) = \tilde{p}(\tilde{\mathbf{s}})$, then the decoupled noise variance $\sigma^2_{t+1}$ of mcB-AMP is equivalent to $\sigma^2_{t+1}$ of cB-AMP in Theorem 2.*

We note that for separable sets defined in Definition 10, the simplification of the complex-valued SE to a real-valued SE follows for mismatched cB-AMP. Theorem 13 allows the MSE for the mismatched prior also to be factored into real-valued SE. This simplification is useful as it reduces the tuning complexity. The separability result from Theorem 13 also extends to that for mcB-AMP in Algorithm 3; the result is shown below in Corollary 18:

**Corollary 18.** *Let the constellation $\mathcal{O}$ be separable. Define $S_{\mathrm{R}} = \mathrm{Re}(S)$ and denote the real-part of $\mathcal{O}$ as $\mathcal{O}^{\mathrm{R}}$. Define $\mathsf{F}^{\mathrm{R}}$ as the mismatched message mean function associated with the mismatched prior. Then, the complex-valued mismatched posterior mean $\mathsf{F}^{\mathrm{mm}}$ can be computed separately in real and imaginary parts as:*

$$\mathsf{F}^{\mathrm{mm}}(S, \tau^2) = \mathsf{F}^{\mathrm{R}}(\mathrm{Re}(S), \tau^2/2) + \mathrm{i}\mathsf{F}^{\mathrm{R}}(\mathrm{Im}(S), \tau^2/2).$$

*Also define the MSE function $\Psi$ for the real-valued prior $S_{\mathrm{R}}$ as:*

$$\Psi^{\mathrm{R}}(\sigma^2, \gamma^2) = \mathbb{E}_{S_{\mathrm{R}}, N_{\mathbb{R}}}\left[\left(\mathsf{F}^{\mathrm{R}}(S_{\mathrm{R}} + \sigma N_{\mathbb{R}}, \gamma^2) - S_{\mathrm{R}}\right)^2\right],$$

*where $N_{\mathbb{R}} \sim \mathcal{N}(0,1)$. Then we have the following relation for $\Psi^{\mathrm{mm}}$ between the complex-valued constellation $\mathcal{O}$ and the real-valued constellation $\mathcal{O}^{\mathrm{R}}$:*

$$\Psi^{\mathrm{mm}}(\sigma^2, \gamma^2) = 2\Psi^{\mathrm{R}}\left(\frac{\sigma^2}{2}, \frac{\gamma^2}{2}\right). \tag{4.8}$$

*Therefore, the recursions in (4.6) and (4.7) can be simplified to:*

$$\gamma_t^2 = \underset{\gamma^2 \geq 0}{\arg\min}\, \Psi^{\mathrm{mm}}\left(\frac{\sigma_t^2}{2}, \frac{\gamma^2}{2}\right),$$
$$\sigma_t^2 = N_0 + \beta\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2) = N_0 + 2\beta\Psi^{\mathrm{R}}\left(\frac{\sigma_t^2}{2}, \frac{\gamma_t^2}{2}\right),$$

Corollary 18 shows that the tuning stage in (4.3) and message mean (4.4) of mcB-AMP algorithm can be computed (often more efficiently) in parallel for real and imaginary dimensions.

### 4.2.3 Optimal Tuning of the Variance Parameter $\tau$

The purpose of (4.3) is to optimally tune the variance parameter $\tau^2$ in every iteration $t$, which is used to compute the posterior mean in (4.5). Before we discuss the tuning procedure in detail, we define we mean by optimally-tuning the variance parameter $\tau^2$. For $t = 1, \ldots, t_{\max}$ iterations, our goal is to minimize the decoupled noise variance $\sigma^2_{t_{\max}+1}$ given by Theorem 16 as the smallest $\sigma^2_{t_{\max}+1}$ minimizes the error probability of our algorithm. To achieve this goal, the optimal choice is to tune the parameters $\hat{\tau}_1^2, \ldots, \hat{\tau}^2_{t_{\max}}$ so that mcB-AMP ultimately leads to the smallest $\sigma^2_{t_{\max}+1}$. We next show that the optimization stage (4.3) which is done separately at every iteration, in fact, achieves the smallest $\sigma^2_{t_{\max}+1}$, i.e., optimally tunes the variance parameters $\hat{\tau}_t^2$. We note that suboptimal choices of $\hat{\tau}_t^2$ can either lead to a higher $\sigma^2_{t_{\max}+1}$ or cause slower converge to the minimal $\sigma^2_{t_{\max}+1}$. We start by defining what we mean by "optimally-tuned" parameters [136].

**Definition 12.** *Assume the large-system limit and denote the decoupled noise variance of mcB-AMP obtained from the sequence $\hat{\tau}_1^2, \ldots, \hat{\tau}^2_{t_{max}}$ as $\sigma^2_{t_{max}+1}(\hat{\tau}_1^2, \ldots, \hat{\tau}^2_{t_{max}})$. A sequence of parameters $\tau_1^{2,\star}, \ldots, \tau^{2,\star}_{t_{max}}$ is optimally-tuned at the iteration $t_{max}$, if and only if for all $(\tau_1^2, \ldots, \tau^2_{t_{max}}) \in [0, \infty)^{t_{max}}$,*

$$\sigma^2_{t_{max}+1}(\tau_1^{2,\star}, \ldots, \tau^{2,\star}_{t_{max}}) \leq \sigma^2_{t_{max}+1}(\tau_1^2, \ldots, \tau^2_{t_{max}}). \tag{4.9}$$

In words, a sequence of optimally-tuned parameters minimizes the decoupled noise variance $\sigma^2_{t_{\max}+1}$ in Theorem 16 given by mcB-AMP after $t_{\max}$ iterations.

The following theorem shows that mcB-AMP leads to the $t_{\max}$ optimally-tuned parameters $\{\tau_1^{2,\star}, \ldots, \tau_{t_{\max}}^{2,\star}\}$. For the sake of brevity, we skip the proof details as it follows the proof in [136, Sec. 4.4] with minor modifications.

**Theorem 19.** *Suppose $\{\tau_1^{2,\star}, \ldots, \tau_{t_{max}}^{2,\star}\}$ are optimally-tuned for iteration $t_{max}$. Then, for any $t < t_{max}$, the parameters $\{\tau_1^{2,\star}, \ldots, \tau_t^{2,\star}\}$ are also optimally-tuned for iteration $t$. Thus, one can obtain $t_{max}$ optimally-tuned variance parameters by optimizing $\tau_1^{2,\star}$ at $t = 1$, and then proceeding iteratively by optimizing $\tau_t^{2,\star}$ until $t = t_{max}$.*

We note that the exact value of the decoupled noise variance $\sigma_t^2$ that is needed for the tuning stage in (4.3) to select $\tau_t^{2,\star}$ is, in general, unknown at iteration $t$. We therefore use an estimate $\hat{\sigma}_t^2 = \frac{1}{B}\left\|\mathbf{r}^t\right\|_2^2$ in step (4.2), which was shown to converge to the true decoupled noise variance $\sigma_t^2$ in the large-system limit [86]. Another implication of Theorem 19 is that the $t_{\max}$ parameters not only achieve the minimum $\sigma_{t_{\max}}^2(\tau_1^{2,\star}, \ldots, \tau_{t_{\max}}^{2,\star})$ under mcB-AMP, but also does so at the fastest convergence rate.

### 4.2.4 Fixed-point Analysis

While the performance of mcB-AMP at every iteration $t = 1, \ldots, t_{\max}$ in the large-system limit can be characterized by the SE recursion equations in Theorem 16, we can analyze the performance of mcB-AMP for $t_{\max} \to \infty$. In this case, the mismatched SE in Theorem 16 converges to the following fixed-point equation:

$$\sigma_\star^2 = N_0 + \beta \min_{\gamma^2 \geq 0} \Psi^{\mathrm{mm}}(\sigma_\star^2, \gamma^2) = N_0 + \beta \Psi_\star(\sigma_\star^2). \tag{4.10}$$

Thus, as $t_{\max} \to \infty$, the decoupled noise variance by mcB-AMP converges to $\sigma_\star^2$ for the relation (4.10). If there are multiple fixed points, we note that mcB-AMP, in

general[1], converges to the largest fixed-point solution to (4.10), which ultimately leads to a higher probability of error than that of the smallest fixed-point solution. In order to provide conditions on the MIMO system to ensure a unique fixed-point solution to (4.10), we use the MRT in Definition 7. By the definition of $\beta_{\mathcal{O}}^{\min}$, for all system ratios $\beta < \beta_{\mathcal{O}}^{\min}$ regardless of the noise variance $N_0$, the fixed-point solution in (4.10) is unique. In Section 4.4, we will use the MRT to analyze when mcB-AMP has a unique fixed-point solution for QAM/PAM constellation sets and establish connections to other convex-based methods.

## 4.3 Mismatched Data Detection with Optimal Tuning: General Case

We now apply the mismatched cB-AMP framework to mismatched data detection in large MIMO systems, and refer to the algorithm as mismatched large MIMO AMP (M-LAMA). In this section, we start by introducing M-LAMA and then, present corresponding M-LAMA for a Gaussian prior.

### 4.3.1 Optimally-Tuned Data Detection with a Gaussian Prior

In this section, we derive a M-LAMA algorithm variant using a mismatched Gaussian prior when the true signals are taken from a constellation set $\mathcal{O}$ with equally likely symbols assuming $\mathbb{E}_{S_0}[|S_0|^2] = E_s$. Without loss of generality, we assume a standard complex Gaussian distribution for the mismatched prior, i.e.,

[1]Convergence to another fixed-point solution is possible if mcB-AMP is initialized sufficiently close to such a fixed point [114].

$\tilde{p}(\tilde{s}_\ell) \sim \mathcal{CN}(0,1)$ as the variance parameter $\tau_t^2$ will be scaled accordingly to $E_s$ in the tuning stage (4.3). For the mismatched Gaussian prior, the message mean function (4.5) is given by $\mathsf{F}^{\mathrm{mm}}(s,\tau^2) = \frac{E_s}{E_s+\tau^2}s$. Using Theorem 16, we can derive the following mismatched SE recursion:

$$\sigma_{t+1}^2 = N_0 + \beta \min_{\gamma^2 \geq 0} \left\{ \frac{E_s^2 \sigma_t^2}{(E_s+\gamma^2)^2} + \frac{E_s \gamma^4}{(E_s+\gamma^2)^2} \right\}, \tag{4.11}$$

which is obtained by substituting $\mathsf{F}^{\mathrm{mm}}(s,\tau^2) = \frac{E_s}{E_s+\tau^2}s$ into (4.7).

Evidently, the mismatched SE recursion (4.11) only depends on the signal energy $E_s$ and no other properties of the true prior $p(\mathbf{s}_0)$. This fact allows us to optimally tune the variance parameters only by knowledge of signal energy $E_s$. Therefore, if the true prior is unknown, but we know the signal energy, one may use M-LAMA to perform data detection. Before we proceed to the fixed-point analysis, the following lemma, with proof in Appendix A.2.2, connects the tuning stage of M-LAMA in (4.3) and (4.11).

**Lemma 20.** *Assume a mismatched Gaussian prior $\tilde{p}(\tilde{s}) \sim \mathcal{CN}(0,1)$ and the power of the true prior is $\mathbb{E}_{S_0}[|S_0|^2] = E_s$. Then, the optimal choice in the tuning stage (4.6) is $\gamma_\star^2 = \sigma_t^2$ which is the global minimizer to (4.11) for a fixed $\sigma_t^2 \geq 0$.*

Thus, the mismatched SE recursion (4.11) reduces to

$$\sigma_{t+1}^2 = N_0 + \beta \frac{E_s}{E_s+\sigma_t^2}\sigma_t^2, \tag{4.12}$$

and the MRT is given by

$$\beta^{\min} = \min_{\sigma^2 \geq 0} \left\{ \left( \frac{\mathrm{d}\Psi^{\mathrm{mm}}(\sigma^2)}{\mathrm{d}\sigma^2} \right)^{-1} \right\} = \min_{\sigma^2 \geq 0} \left( 1 + \frac{\sigma^2}{E_s} \right)^2 = 1.$$

Therefore M-LAMA has a unique fixed point when $\beta \leq 1$ regardless of the noise variance $N_0$. Interestingly, the fixed-point equation of (4.11) of this algorithm

corresponds to the decoupled noise variance given by the linear MMSE detector [13, 76]. If we define signal-to-interference ratio (SIR) as SIR $= 1/\sigma^2$, and let $t_{\max} \to \infty$, the fixed-point solution of (4.12) coincides to the SIR given by the linear MMSE detector in the large-system limit [13, 76, 128]. Hence, for a mismatched Gaussian prior, M-LAMA achieves the performance of the linear MMSE detector. We note that the proofs given in [13, 76, 128] use results from random matrix theory, whereas our analysis uses the mismatched SE framework proposed in Theorem 16. Furthermore, our result is constructive, i.e., M-LAMA is a computationally efficient algorithm that implements linear MMSE detection without the need of computing a matrix inverse.

### 4.3.2 Suboptimal Data Detection with a Gaussian Prior

We can replace the optimal tuning stage (4.6) by a fixed (and predetermined) variance parameter choice for $\gamma_t^2$, which leads to a suboptimal, mismatched algorithm, referred to as suboptimal M-LAMA (short SM-LAMA). We now show that this approach leads to other well-known linear data detectors. In particular, by considering the following two variance parameter choices $\gamma_t^2 \to 0$ and $\gamma_t^2 \to \infty$ in (4.6), we obtain the following mismatched SE recursions:

$$\text{(ZF)} \quad \sigma_{t+1}^2 = N_0 + \beta \lim_{\gamma_t^2 \to 0} \Psi^{\text{mm}}(\sigma_t^2, \gamma_t^2) = N_0 + \beta \sigma_t^2,$$

$$\text{(MF)} \quad \sigma_{t+1}^2 = N_0 + \beta \lim_{\gamma_t^2 \to \infty} \Psi^{\text{mm}}(\sigma_t^2, \gamma_t^2) = N_0 + \beta \, \mathbb{V}\text{ar}_{S_0}[S_0],$$

respectively.

We note that (ZF) has a unique fixed point when $\beta < 1$ regardless of the noise variance $N_0$, whereas (MF) has unique fixed point regardless for any $\beta$. The

solution to the fixed-point equation (ZF) when $\beta < 1$ and (MF) coincides exactly to the SIR given by ZF and MF detector in the large-system limit [13,76,128,137], respectively. Hence, suboptimal variance parameter choices for $\gamma_t^2$ leads SM-LAMA to achieve the performance of ZF and MF detector.

## 4.4 Mismatched Data Detection with Optimal Tuning: QAM/PAM Constellation

In this section, we present a M-LAMA algorithm variant that improves upon M-LAMA for a Gaussian prior presented in Section 4.3 for QAM and PAM constellations, which are frequently used in practical wireless systems. We will do so by selecting a prior that closely resembles the true prior, unlike Section 4.3, where we assumed a Gaussian prior and only used the signal energy $E_s$. We will present two algorithms of M-LAMA that assumes a (i) uniform hypercube prior and (ii) Gray coding based approximation. For each algorithm variant, we describe the optimal tuning procedure, and present an example of a suboptimal method that avoids the tuning completely. Throughout the rest of this section, we will assume that the true prior is from a QAM or PAM constellation set $\mathcal{O}$ with equally likely symbols.

### 4.4.1 Optimally-Tuned Data Detection with a Uniform Hypercube Prior

In this section, we derive a M-LAMA algorithm variant using a mismatched uniform distribution within a hypercube with length $2\alpha$ around the true prior distribution of square constellations (e.g., QPSK and QAM) . For example, for a QPSK system with equally likely symbols, we use a mismatched prior that is distributed uniformly in the interval $[-1,+1]$ for both the real and imaginary part, so $\alpha = 1$. For this mismatched prior, we use Corollary 18 to compute the posterior mean function independently for the real and imaginary part; the posterior mean function $\mathsf{F}^{\mathrm{mm}}$ and its first derivative are given by:

$$\begin{aligned}\mathsf{F}^{\mathrm{mm}}(\hat{s},\tau^2) &= \hat{s}^{\mathrm{R}} + \frac{\tau^2}{2}\nu_-(\hat{s}^{\mathrm{R}},\tau^2/2) \\ &\quad + \mathrm{i}\left(\hat{s}^{\mathrm{I}} + \frac{\tau^2}{2}\nu_-(\hat{s}^{\mathrm{I}},\tau^2/2)\right) \qquad (4.13)\\ \mathsf{F}'^{\mathrm{mm}}(\hat{s},\tau^2) &= 1 - \frac{1}{2}\left(\hat{s}^{\mathrm{R}}\nu_-(\hat{s}^{\mathrm{R}},\tau^2/2) + \alpha\nu_-(\hat{s}^{\mathrm{R}},\tau^2/2)\right) \\ &\quad - \frac{1}{2}\left(\hat{s}^{\mathrm{I}}\nu_-(\hat{s}^{\mathrm{I}},\tau^2/2) + \alpha\nu_-(\hat{s}^{\mathrm{I}},\tau^2/2)\right) \\ &\quad - \frac{\tau^2}{4}\left(\nu_-^2(\hat{s}^{\mathrm{R}},\tau^2/2) + \nu_-^2(\hat{s}^{\mathrm{I}},\tau^2/2)\right), \qquad (4.14)\end{aligned}$$

where we use the following shorthand notations: $\hat{s}^{\mathrm{R}} = \mathrm{Re}(\hat{s})$, $\hat{s}^{\mathrm{I}} = \mathrm{Im}(\hat{s})$, and we define the following functions:

$$\begin{aligned}\nu_-(\hat{s},\tau^2) &= \frac{\exp\left(-\frac{1}{2\tau^2}(\hat{s}+\alpha)^2\right) - \exp\left(-\frac{1}{2\tau^2}(\hat{s}-\alpha)^2\right)}{\sqrt{2\pi\tau^2}\left(\Phi\left(\frac{\hat{s}+\alpha}{\tau}\right) - \Phi\left(\frac{\hat{s}-\alpha}{\tau}\right)\right)},\\ \nu_+(\hat{s},\tau^2) &= \frac{\exp\left(-\frac{1}{2\tau^2}(\hat{s}+\alpha)^2\right) + \exp\left(-\frac{1}{2\tau^2}(\hat{s}-\alpha)^2\right)}{\sqrt{2\pi\tau^2}\left(\Phi\left(\frac{\hat{s}+\alpha}{\tau}\right) - \Phi\left(\frac{\hat{s}-\alpha}{\tau}\right)\right)}.\end{aligned}$$

The mismatched SE recursion is obtained by Theorem 16 and can be evaluated numerically.

Before we proceed, we discuss two main disadvantages of M-LAMA algorithm for practical systems: (i) The M-LAMA algorithm under hypercube prior is not particularly efficient from a hardware perspective as the function in (4.13) involves transcendental functions. In fact, this algorithm requires the evaluation of the functions $\nu_+(\hat{s}, \tau^2)$ and $\nu_-(\hat{s}, \tau^2)$ in every iteration and for every antenna, which require—similar to that of the optimal LAMA algorithm [39]—excessively high numerical precision (see, e.g., [138]) and a large dynamic range. (ii) Furthermore, the tuning stage (4.3) turns out to be non-trivial—while a grid search or bisection methods are viable numerical methods to find a minimum, corresponding hardware designs are impractical.

### 4.4.2 Suboptimal Data Detection with a Uniform Hypercube Prior

Analogously to the ZF detector in Section 4.3.2, we can derive a sub-optimal variant of M-LAMA (SM-LAMA) with the uniform hypercube prior from Section 4.4.1, where we replace the tuning stage in (4.3) by the fixed choice $\tau^t \to 0$. This choice leads to a much simpler algorithm compared to the optimally-tuned M-LAMA algorithm and thus, makes the performance analysis more accessible. Firstly, the posterior mean function simplifies to:

$$\lim_{\tau^2 \to 0} \mathsf{F}^{\mathrm{mm}}(\hat{s}, \tau^2) = \hat{s}^{\mathrm{R}} + \operatorname{sign}(\hat{s}^{\mathrm{R}}) \min\Big\{\alpha - |\hat{s}^{\mathrm{R}}|, 0\Big\} + \mathrm{i}\Big(\hat{s}^{\mathrm{I}} + \operatorname{sign}(\hat{s}^{\mathrm{I}}) \min\Big\{\alpha - |\hat{s}^{\mathrm{I}}|, 0\Big\}\Big),$$

which can be evaluated efficiently. Secondly, computing $\lim_{\tau^2 \to 0} \mathsf{F}'^{\mathrm{mm}}(\hat{s}, \tau^2)$ is straightforward, which is given by $\lim_{\tau^2 \to 0} \mathsf{F}'^{\mathrm{mm}}(\hat{s}, \tau^2) = \frac{1}{2}\mathbb{I}(|\hat{s}^{\mathrm{R}}| < \alpha) + \frac{1}{2}\mathbb{I}(|\hat{s}^{\mathrm{I}}| < \alpha)$. Lastly, by letting $\tau^2 \to 0$, the tuning stages in (4.2) and (4.3) are

unnecessary as we are fixing the variance parameter $\tau^2$. Since $\mathsf{F}^{\mathrm{mm}}$ is composed of piece-wise linear functions, we can explicitly state the mismatched SE recursion in (4.7) for SM-LAMA for $M^2$-QAM constellations with the aid of Corollary 18. The result for $M$-PAM is shown below in Lemma 21 with derivation in Appendix A.2.3.

**Lemma 21.** *Assume a M-PAM constellation on a real-valued system. Suppose we run SM-LAMA with letting $\hat{\tau}_t^2 \to 0$ for every iteration $t$. Then, the mismatched SE recursion for SM-LAMA is given by $\sigma_t^2 = N_0 + \beta \Psi_{\mathrm{PAM}}^{\mathrm{mm}}(\sigma_t^2)$, where*

$$\Psi_{\mathrm{PAM}}^{\mathrm{mm}}(\sigma^2) = \sigma^2 + \frac{2}{M}\sum_{k=1}^{M/2}\left[(\bar{\alpha}_k^2 - \sigma^2)Q\left(\frac{\bar{\alpha}_k}{\sigma}\right) + (\alpha_k^2 - \sigma^2)Q\left(\frac{\alpha_k}{\sigma}\right) - \frac{\sigma}{\sqrt{2\pi}}\bar{\alpha}_k \exp\left(-\frac{\bar{\alpha}_k^2}{2\sigma^2}\right) - \frac{\sigma}{\sqrt{2\pi}}\alpha_k \exp\left(-\frac{\alpha_k^2}{2\sigma^2}\right)\right],$$

*where $\alpha = M-1$, $\bar{\alpha}_k = \alpha - (2k-1)$, $\alpha_k = \alpha + (2k-1)$.*

Now that we have the mismatched SE recursion to $M$-PAM, we can use the result to obtain SE recursion for $M^2$-QAM by using Corollary 18.

**Corollary 22.** *For a $M^2$-QAM constellation with SM-LAMA for $\hat{\tau}_t^2 \to 0$, the mismatched SE recursion is given by:*

$$\begin{aligned}\sigma_{t+1}^2 &= N_0 + \beta \Psi_{\mathrm{QAM}}^{\mathrm{mm}}(\sigma_t^2) \\ &= N_0 + \beta\frac{2}{M}\sum_{k=1}^{M/2}\left[\sigma_t^2 + \left(2\bar{\alpha}_k^2 - \sigma_t^2\right)Q\left(\frac{\bar{\alpha}_k}{\sigma_t/\sqrt{2}}\right) - \frac{\sigma_t}{\sqrt{\pi}}\left(\bar{\alpha}_k \exp\left(-\frac{\bar{\alpha}_k^2}{\sigma_t^2}\right) + \alpha_k \exp\left(-\frac{\alpha_k^2}{\sigma_t^2}\right)\right) + \left(2\alpha_k^2 - \sigma_t^2\right)Q\left(\frac{\alpha_k}{\sigma_t/\sqrt{2}}\right)\right], \end{aligned} \tag{4.15}$$

Now that we have summarized the mismatched SE recursion for both $M$-PAM and $M^2$-PAM for SM-LAMA, we now present conditions on the system

ratio $\beta$ where SM-LAMA has a unique fixed point. The following Lemma 23, with proof in Appendix A.2.4, shows that the MRT of SM-LAMA for $M^2$-QAM is given by $\beta_{\mathcal{O}}^{\min} = (1 - 1/M)^{-1}$.

**Lemma 23.** *Assume that $S_0$ follows from $M^2$-QAM constellation with equally likely symbols. Then the minimum recovery threshold (MRT) for SM-LAMA is given by $\beta_{M^2\text{-QAM}}^{\min} = (1 - 1/M)^{-1}$.*

We also obtain the same MRT for a $M$-PAM constellation by Lemma 14.

**Corollary 24** (SM-LAMA has the same MRT for $M^2$-QAM and $M$-PAM (under a real-valued system))**.**

We now show that this SM-LAMA variant achieves the same performance as a well-known relaxation of the maximum likelihood data detection problem [15, 131–133]. This algorithm, known as box relaxation (BOX) detector, solves the following convex problem [126]:

$$\hat{\mathbf{s}} = \underset{\tilde{\mathbf{s}} \in \mathbb{C}^U}{\arg\min} \ \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{s}}\|_2 \quad \text{subject to } \|\tilde{\mathbf{s}}\|_\infty \leq \alpha \tag{4.16}$$

and slices the individual entries of $\hat{\mathbf{s}}$ onto the $M^2$-QAM (or $M$-PAM) constellations. The next result shows that SM-LAMA achieves the same error-rate performance as the BOX detector in (4.16), while providing a simple and computationally efficient algorithm. The proof is given in Appendix A.2.5.

**Lemma 25.** *For a complex-valued MIMO system with $M^2$-QAM constellations (or a real-valued MIMO system with M-PAM constellations), SM-LAMA achieves the same performance as convex relaxed methods in [14].*

We emphasize that in [14], the results were only pertained to real-valued system, whereas our analysis connect both complex-valued $M^2$-QAM and real-valued $M$-PAM system. The case for QPSK (BPSK for real-valued system) was

shown in [15]. In addition, our analysis is constructive; SM-LAMA algorithm variant readily works for both QAM and PAM constellations.

### 4.4.3 Optimally-Tuned Data Detection with Gray Coding Based Approximation

In this section, we derive M-LAMA algorithm variant by exploiting Gray coding that are used for practical systems that use QAM constellation. For the sake of brevity, we will derive the posterior mean function and analysis for 16-QAM and uniform priors only; our approach can be easily extended to higher order QAM constellations and non-uniform priors. We first start by noting that (3.19) for 16-QAM for uniform priors corresponds to

$$\mathsf{F}_{\text{16-QAM}}(\hat{s}, \tau^2) = \mathsf{F}_{\text{4-PAM}}(\hat{s}^{\text{R}}, \tau^2/2) + \mathrm{i}\mathsf{F}_{\text{4-PAM}}(\hat{s}^{\text{I}}, \tau^2/2),$$

where we utilized the separable property of 16-QAM as described in Corollary 18. Here, we introduced a shorthand notation for the posterior mean for 4-PAM, which is:

$$\begin{aligned}\mathsf{F}^{\text{4-PAM}}(\hat{s}, \tau^2) &= \frac{-3e_{-3} - 1e_{-1} + e_1 + 3e_3}{e_{-3} + e_{-1} + e_1 + e_3}, \\ e_a &= \exp\left(-\frac{1}{2\tau^2}(\hat{s} - a)^2)\right).\end{aligned}$$

We note that the above $\mathsf{F}^{\text{4-PAM}}$ function can be rewritten as:

$$\mathsf{F}^{\text{4-PAM}}(\hat{s}, \tau^2) = \sum_{a \in \mathcal{O}^{\text{4-PAM}}} w_a(\hat{s}, \tau^2) a,$$

where $\mathcal{O}^{\text{4-PAM}} = \{\pm 3, \pm 1\}$ and $w_a(\hat{s}, \tau^2)$ is a weight distribution so that

$$\sum_{a \in \mathcal{O}^{\text{4-PAM}}} w_a(\hat{s}, \tau^2) = 1.$$

For the 4-PAM example above, we have

$$w_a(\hat{s}, \tau^2) = \frac{e_a}{e_{-3} + e_{-1} + e_1 + e_3} \tag{4.17}$$

The high arithmetic precision requirement of computing the posterior mean mainly stems from the computation of (4.17); this is due to the fact that $e_a$ decays exponentially fast to zero for small $\tau^2$. Thus, the computation of (4.17) requiring high numerical precision, makes the design of efficient VLSI architectures challenging.

We now propose an approximation of (4.17) that not only relieves the arithmetic precision requirements, but also achieves greater performance than that given by a hypercube prior presented in Section 4.4.1. To do so, we exploit Gray coding [57], which are commonly used in practical wireless standards.

Assuming that each bit-wise entry is independent, we exploit Gray coding to decompose each entry of the weight distribution as:

$$w_{-3} = (1 - p_1)(1 - p_0), \quad w_{-1} = (1 - p_1)p_0,$$
$$w_1 = p_1 p_0, \quad w_3 = p_1(1 - p_0),$$

where we dropped the index $(\hat{s}, \tau^2)$ for simplicity. Here, we introduced $p_b$, $b = 0, 1$ that represent the probability that $b$th bit is equal to 1. Under Gray coding and independence bit assumption, it is easy to see that $p_1 = w_1 + w_3$ and $p_0 = w_{-1} + w_1$. Thus, with the new notation, $\mathsf{F}^{\text{Gray}}_{\text{4-PAM}}(\hat{s}, \tau^2) = (1 - 2p_1)(2p_0 - 3)$.

Now, instead of computing $p_1$ and $p_0$ directly, we compute the log-likelihood ratio (LLR) $\Lambda_b = \log \frac{p_b}{1-p_b}$ and use the relation $p_b = \frac{1}{2}(1 + \tanh(\frac{1}{2}\Lambda_b))$ for bits $b = 0, 1$, which can be computed efficiently. With this formulation, the derivative

becomes:

$$\begin{aligned}\mathsf{F}'^{\text{Gray}}_{\text{4-PAM}}(\hat{s},\tau^2) &= (1-\tanh^2(\bar{\Lambda}_1))(2-\tanh(\bar{\Lambda}_0))\bar{\Lambda}_1' \\ &\quad -\tanh(\bar{\Lambda}_1)(1-\tanh^2(\bar{\Lambda}_0))\bar{\Lambda}_0' \\ &= 4p_1(1-p_1)(3-2p_0)\bar{\Lambda}_1' + 4(1-2p_1)p_0(1-p_0)\bar{\Lambda}_0',\end{aligned}$$

where we introduced a shorthand notation $\bar{\Lambda}_b = \Lambda_b/2$ for both $b = 0, 1$. We note that the above expression of $\mathsf{F}'$ can be implemented efficiently in hardware as $1-\tanh^2(\cdot)$ has low dynamic range. Now that we have described the procedure for computing the posterior mean function, we will describe how to compute the log-likelihood ratios $\Lambda_0$ and $\Lambda_1$.

**Computation of the Log-Likelihood Ratio**

In this section, we describe computationally efficient methods to compute the LLRs $\Lambda_0$ and $\Lambda_1$. As described earlier, the following section assumes 16-QAM constellation; the methodology presented here can easily be extended to higher order QAM constellations. We propose two methods to compute $\Lambda_0$ and $\Lambda_1$: (i) direct approach and (ii) low-complexity approach via max-log approximation.

(i) The direct computation of the LLR is straightforward and is computed by noting the fact that $p_1 = w_1 + w_3$ and $p_0 = w_{-1} + w_1$.

$$\begin{aligned}\Lambda_0 &= \log\frac{p_0}{1-p_0} = \log\frac{w_{-1}+w_1}{w_{-3}+w_3} \\ &= 8\rho + \log\frac{\cosh(2\rho s)}{\cosh(6\rho s)}, \\ \Lambda_1 &= 8\rho s + \log\frac{\cosh(2\rho(s-2))}{\cosh(2\rho(s+2))},\end{aligned}$$

where we defined shorthand notation $\rho = 1/\tau^2$.

(ii) Although the expressions for $\Lambda_0$ and $\Lambda_1$ are straightforward, computing them can be quite challenging to the the transcendental nature of the ratio term of cosh functions. Thus, to reduce the dynamic range for the computation of the LLRs, we propose a low-complexity method to approximate the exact LLRs $\Lambda_0$ and $\Lambda_1$ via max-log approximation [139]. The idea here is to notice that $\log(\exp(a)+\exp(-a)) = |a| + \log(1+\exp(-2|a|)) \simeq a$ for large values of $a$ as $\exp(-2|a|)$ quickly converges to 0. Thus, applying the max-log approximation to $\Lambda_0$ results in:

$$\begin{aligned}
\log\frac{\cosh(2\rho s)}{\cosh(6\rho s)} &= \log\cosh(2\rho s) - \log\cosh(6\rho s) \\
&= 2\rho|s| + \log(1+\exp(-4\rho|s|)) \\
&\quad - 6\rho|s| - \log(1+\exp(-12\rho|s|)) \\
&\overset{(a)}{\simeq} 2\rho|s| - 6\rho|s| = -4\rho|s|.
\end{aligned}$$

Similarly for $\Lambda_1$ we have,

$$\log\frac{\cosh(2\rho(s-2))}{\cosh(2\rho(s+2))} \overset{(b)}{\simeq} 2\rho|s-2| - 2\rho|s+2|.$$

We note that in $(a)$ and $(b)$, we have ignored the correction term $\log(1+\exp(-c|s|))$; one may include an approximated value of the correction term to mitigate the loss of max-log approximation at a small overhead in complexity. However, we will show in Section 4.5 that the proposed max-log approximation does not result in any loss in error-rate performance. Therefore, we have the

LLRs $\Lambda_0$ and $\Lambda_1$ from the max-log approximation as:

$$\begin{aligned}\Lambda_0^{\text{max-log}} &= 4\rho(2-|s|),\\ \Lambda_1^{\text{max-log}} &= 2\rho(4s+|s-2|-|s+2|)\\ &= \begin{cases} 8\rho(s+1) & s<-2,\\ 4\rho s & -2\le s<2,\\ 8\rho(s-1) & s>2.\end{cases}\end{aligned}$$

We note that the resulting LLRs can be computed very efficiently without any special functions and requires low precision range compared to the exact LLRs obtained directly.

### 4.4.4 Suboptimal Data Detection with Gray Coding Based Approximation

As described in Section 4.2.3, optimal tuning is necessary with a mismatched prior to mitigate the performance loss. Optimal tuning for Gray-coding based approximation involves solving for the optimization problem in (4.6). For the Gray-coding based approximation, we have the posterior mean with the exact LLR computation as:

$$\begin{aligned}\mathsf{F}_{\text{4-PAM}}^{\text{Gray}}(\hat{s},\tau^2) &= \tanh(\bar{\Lambda}_1)(2-\tanh(\bar{\Lambda}_0))\\ &= \tanh\left(\frac{4}{\tau^2}\hat{s}+\frac{1}{2}\log\frac{\cosh\left(\frac{2}{\tau^2}(\hat{s}-2)\right)}{\cosh\left(\frac{2}{\tau^2}(\hat{s}+2)\right)}\right)\\ &\quad\times\left(2-\tanh\left(\frac{4}{\tau^2}+\frac{1}{2}\log\frac{\cosh\left(\frac{2}{\tau^2}\hat{s}\right)}{\cosh\left(\frac{6}{\tau^2}\hat{s}\right)}\right)\right),\end{aligned}$$

and the posterior mean with max-log approximation as:

$$\mathsf{F}_{\text{4-PAM}}^{\text{max-log}}(\hat{s},\tau^2) = \tanh\left(\frac{1}{\tau^2}(4\hat{s}+|\hat{s}-2|-|\hat{s}+2|)\right) \times \left(2-\tanh\left(\frac{2}{\tau^2}(2-|\hat{s}|)\right)\right).$$

Solving for the optimization problem in (4.6) with the posterior mean functions above, however, is also non-trivial. Similar to optimal tuning with a uniform hypercube prior in Section 4.4.1, grid search or bisection methods are viable, but are impractical as it requires additional computational overhead. Based on numerical calculations for optimization of (4.6) for 16-QAM constellation, we observe that letting $\hat{\tau}^2 = \sigma^2$ provides a good estimate that performs very close to the optimally-tuned value $\tau^2$ without any additional overhead.

## 4.5 Numerical Results

We now compare the error-rate performance of the proposed M-LAMA algorithm variants. The mismatched SE framework Theorem 16 enables an exact error-rate analysis in the large-system-limit, which does not require numerical simulations. In Figure 4.1 and 4.2, we compare the performance of our M-LAMA algorithms in the large system limit to that of the performance in a $256 \times 128$ finite-dimensional system MIMO system for two constellations, QPSK and 16-QAM. The error-rate performance predicted by the SE framework in the large-system limit are shown in either solid, dashed, or dotted lines, whereas the error-rate performance obtained by numerical simulations are shown in markers. For both simulations, we also compare our results to LAMA, which was shown to achieve near-IO performance in finite systems and IO performance in the large-system limit for $\beta = 0.5$ for both constellations.

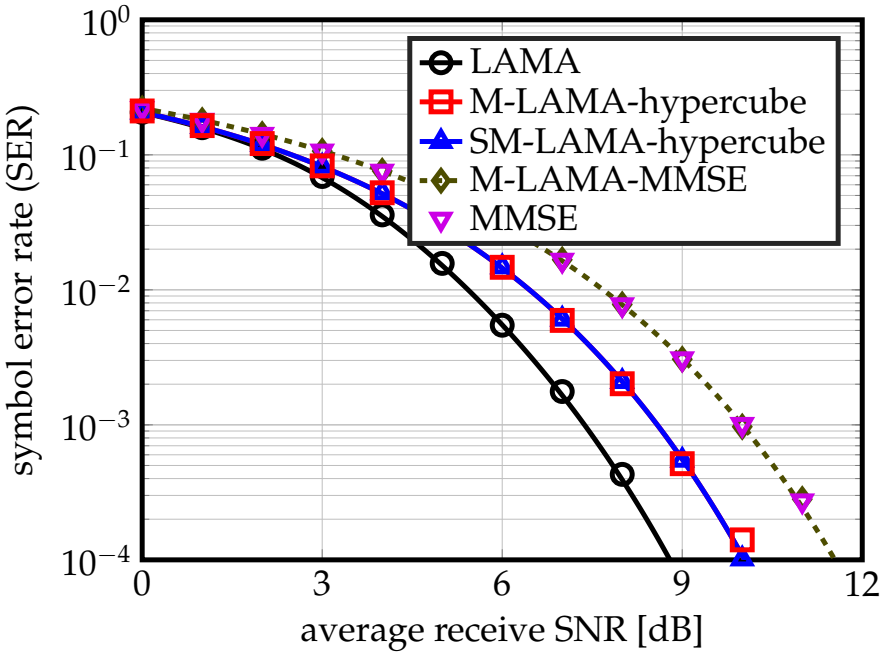



Figure 4.1: Symbol error rate for M-LAMA algorithm and its variants for $128 \times 64$ MIMO system with QPSK constellation and 10 algorithm iterations. The error-rate performance predicted by the SE framework in the large-system limit are shown in lines whereas error-rate performance obtained by numerical simulations are shown in markers. For QPSK, SM-LAMA under uniform hypercube prior performs within 1 dB from LAMA, which achieves IO performance in the large-system limit.

Fig. 4.1 shows the error-rate performance of M-LAMA algorithms for QPSK constellation. We show the performance of three mismatched LAMA algorithms, optimally-tuned M-LAMA for a Gaussian prior (M-LAMA-MMSE), and optimally- and sub-optimal tuned M-LAMA for hypercube prior; we include error-rate performance of LAMA as a baseline. We exclude Gray-coding based approximation as it has no performance loss for QPSK constellation; hence, for QPSK, Gray-coding based approximation achieves identical error-rate performance to that given by LAMA. We first observe that M-LAMA algorithms for a finite dimension achieve similar error-rate performance to its error-rate performance in the large-system limit. In particular, M-LAMA for optimally-tuned Gaussian prior (M-LAMA-MMSE) achieves near-identical error-rate performance

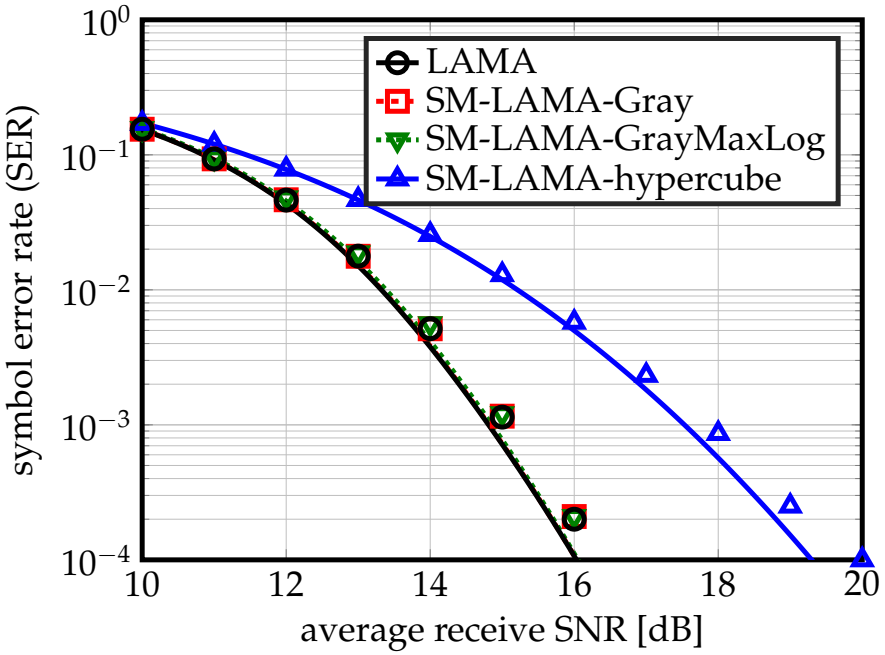



Figure 4.2: Symbol error rate for M-LAMA algorithm and its variants for $128 \times 64$ MIMO system with 16-QAM and 5 algorithm iterations. The error-rate performance predicted by the SE framework in the large-system limit are shown in lines whereas error-rate performance obtained by numerical simulations are shown in markers. For 16-QAM, SM-LAMA with Gray-coding based approximation performs as good as LAMA. Thus, M-LAMA can achieve near-IO optimal error-rate performance at a lower computational complexity by using intentionally selected mismatched priors.

to actual linear MMSE detection, which agrees with SE analysis. In particular, the optimally-tuned M-LAMA and its ZF variant SM-LAMA, for uniform hypercube prior performs within 1 db of the LAMA algorithm, which was shown to achieve IO performance in the large-system limit.

Fig. 4.2 shows the error-rate performance of M-LAMA for 16-QAM constellation. We provide error-rate performance of both Gray-coding based approximations, and sub-optimally tuned M-LAMA for hypercube prior. Compared to Fig. 4.1, there exists performance gap between asymptotic predictions by SE and numerical simulations; the gap decreases for larger systems. We observe that Gray-coding based approximation provides excellent error-rate performance

compared to that given by LAMA. Moreover, there does not exist any performance loss of using max-log approximation for Gray coding based approximation compared that without max-log approximation. Fig. 4.2 reiterates the fact that M-LAMA with carefully selected intentional mismatched priors can achieve near-optimal error-rate performance at a lower computational complexity and without the need for complicated transcendental functions.

## 4.6 Conclusions

We have presented the M-LAMA algorithm along with mismatched SE recursion. We have shown that for a mismatched Gaussian prior, M-LAMA and suboptimal SM-LAMAs achieve the same performance as linear MMSE, and ZF and MF detector. For a QAM constellation, we presented two variants of M-LAMA and characterized the SE performance for a uniform hypercube prior and Gray coding based approximation. For a mismatched uniform hypercube prior, we established conditions on the system ratio $\beta$ for which SM-LAMA has a unique fixed point under $M^2$-QAM constellation. Although the presented results are exclusively for the large-system limit, our simulations indicate that M-LAMA and its variants achieve near-IO performance in realistic, finite dimensional systems.

CHAPTER 5

# DECENTRALIZED DATA DETECTION IN MASSIVE MU-MIMO

## 5.1 Introduction

Massive multi-user (MU) multiple-input multiple-output (MIMO) is widely believed to be a key technology for next-generation wireless systems [30, 53, 140]. By equipping the infrastructure base-stations (BSs) with hundreds or thousands of active antenna elements and serving tens or hundreds of user equipments (UEs) simultaneously and in the same frequency band, massive MU-MIMO promises orders-of-magnitude improvements in spectral efficiency and energy efficiency compared to traditional, small-scale MIMO [5, 6]. However, the large number of antennas at the BS causes significant challenges when implementing this technology. One of the most prominent challenges is the excessively high amount of fronthaul data that must be transferred from the radio-frequency (RF) antenna units at the BS antenna array to the baseband processing unit (BBU) [16–19]. For example, the fronthaul data rates (from RF chains to the BBU) exceed 200 Gbit/s for a massive MU-MIMO system with 128 BS antennas, each using two 10 bit analog-to-digital converters (for in-phase and quadrature components) operating at 80 MS/s sampling rate. Such high data rates not only exceed the bandwidth of existing high-speed interconnect standards, such as the common public radio interface (CPRI) [20], but will also approach the limits of existing chip input/output (I/O) interfaces in terms of bandwidth and power dissipation [21]. Furthermore, traditional data-detection algorithms that achieve near-optimal spectral efficiency in the MU-MIMO uplink [6], such as zero-forcing (ZF) and linear minimum mean-square error (L-MMSE)-based

equalization, rely on centralized processing in a single computing fabric, which results in excessively high complexity for large antenna arrays [17, 22].

### 5.1.1 Decentralized Baseband Processing

In order to mitigate the bandwidth and computing bottlenecks of centralized massive MU-MIMO architectures, existing testbeds either distribute the most critical baseband processing tasks in the frequency domain or use matched filtering (MF). Concretely, the testbeds described in [141–144] parallelize the key baseband processing tasks across the subcarriers of orthogonal frequency-division multiplexing (OFDM)-based systems. While this approach enables high parallelism, it requires that each frequency cluster obtains data from *all* BS antennas, which alone does not enable one to scale such systems to thousands of antenna elements [17]. In contrast to frequency parallelization, MF enables antenna parallelization that divides array into independent clusters [53, 145]; this approach significantly reduces the interconnect bandwidth between the RF chains and the BBUs. MF, however, suffers from low spectral efficiency for realistic antenna configurations and high-rate modulation and coding schemes [6]. Consequently, realizing massive MU-MIMO in practice requires solutions that reduce the interconnect and chip I/O bandwidth as well as the baseband processing complexity per computing fabric, without sacrificing spectral efficiency.

Decentralized baseband processing (DBP) has been proposed in [17] to alleviate the fronthaul and I/O bandwidth bottlenecks, and enables parallel baseband processing across BS antennas on multiple computing fabrics, such as application-specific integrated circuits (ASICs), field-programmable gate arrays (FPGAs), or

graphics processing units (GPUs) [146,147], while achieving high spectral efficiency. The idea of DBP is to partition the BS antenna array into $C$ independent clusters, each associated with local computing fabrics that carry out the necessary RF and baseband processing tasks in a decentralized and parallel fashion. The algorithms proposed in [17] perform linear equalization and precoding in an iterative manner by exchanging consensus information among the clusters. However, implementation results on a GPU cluster revealed that the transfer latency of such consensus-sharing methods are limiting the achievable throughput. To avoid this drawback, references [16,21,148,149] recently proposed *feedforward* architectures that minimize the transfer latency.

### 5.1.2 Relevant Prior Art

Decentralized baseband processing (DBP) for massive MI-MIMO systems has been proposed in [17] together with consensus-sharing equalization and precoding algorithms. Distributed processing across antenna elements is also a critical component of coordinated multipoint (CoMP) [150] and cloud radio access networks (CRANs) [151] for multi-cell transmission. While all these architectures and algorithms are able to reduce the raw baseband data rates and mitigate the computation bottlenecks, their performance has not been analyzed and the achievable throughput suffers from high interconnect latency caused by iterative exchange of consensus information. To avoid the consensus sharing among clusters, we focus on decentralized *feedforward* architectures that minimize the transfer latency and enable a theoretical performance analysis.

Feedforward architectures for decentralized massive MU-MIMO equaliza-

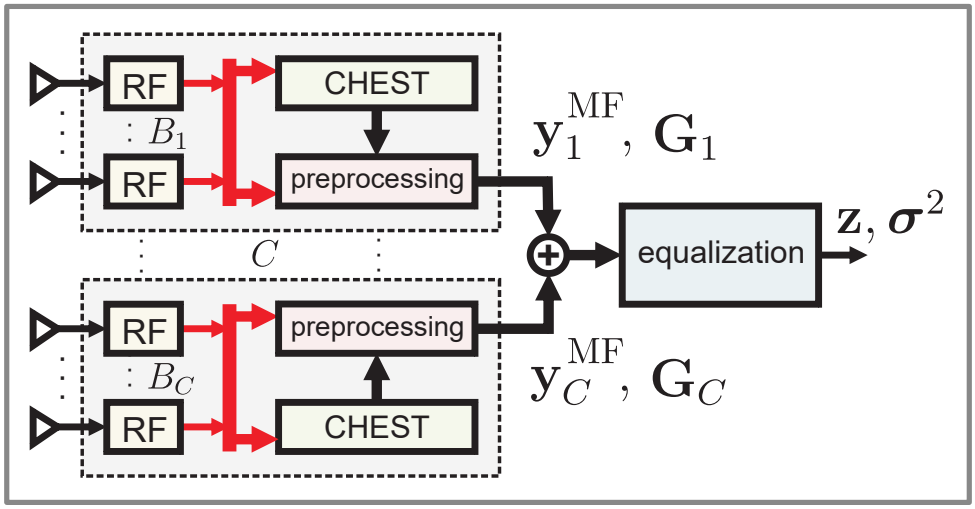
tion have been proposed in [16, 21, 148, 149]. The present chapter extends our theoretical results from [149] and, in contrast to [16, 21, 148], provides two distinct architectures and a corresponding SINR analysis for a range of linear and nonlinear equalization algorithms.

The post-equalization SINR performance of *centralized* linear equalization algorithms, such as MF, ZF, and L-MMSE, has been analyzed in [13, 76, 128, 152] in the large-system limit. We will investigate the SINR performance of these algorithms for the two proposed decentralized feedforward architectures, and also investigate the efficacy of nonlinear equalization for decentralized massive MU-MIMO architectures.

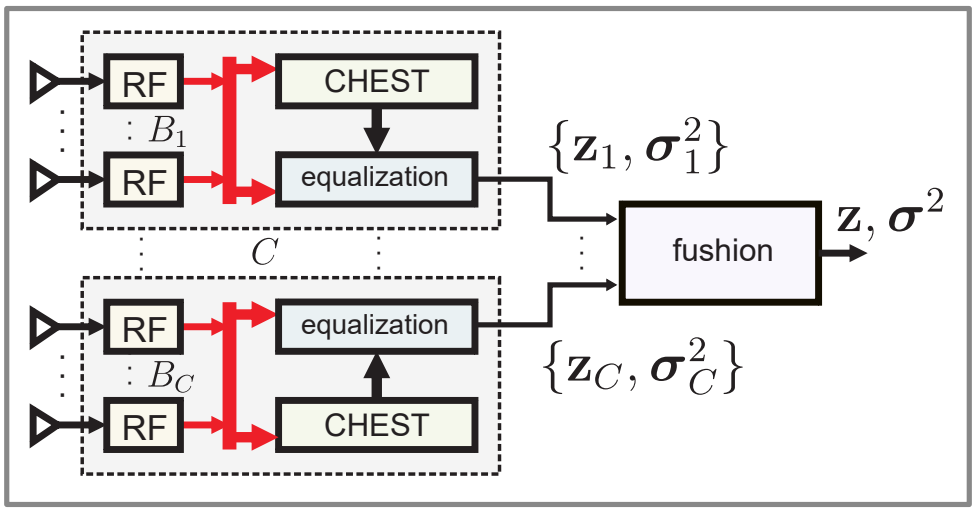
Nonlinear equalization for massive MU-MIMO systems via approximate message passing (AMP) has been studied in [38, 39, 103, 153]. A distributed version of AMP has been proposed in [154] for compressive sensing applications [155, 156]. The key differences of our nonlinear equalization algorithm to these results are as follows: (i) We consider decentralized feedforward architectures; (ii) the methods in [38, 39, 103, 153] are centralized; (iii) the distributed AMP-based method in [154] requires iterative consensus sharing; (iv) we analyze the post-equalization SINR and error-rate performance in massive MU-MIMO systems.

## 5.2 Decentralized Equalization Architectures

We start by introducing the considered massive MU-MIMO system model and the basics of equalization-based data detection. We then discuss the two feedforward equalization architectures for DBP depicted in Fig. 5.1, and detail the SINR




(a) Partially decentralized (PD) equalization architecture.




(b) Fully decentralized (FD) equalization architecture.

Figure 5.1: Partially decentralized (PD) and fully decentralized (FD) feedforward equalization architectures for the massive MU-MIMO uplink. The antenna array is divided in $C$ clusters, each associated with local radio-frequency (RF) processing and channel estimation (CHEST). (a) PD performs decentralized CHEST and preprocessing; equalization is performed in a centralized fashion and operates on a low-dimensional data (dimension is the number of UEs). (b) FD performs CHEST, preprocessing, and equalization in a decentralized manner; the final equalization result is formed by a weighted average of local estimates.

analysis framework that we will use throughout this section.

### 5.2.1 Uplink System Model and Equalization

We consider a narrowband massive MU-MIMO uplink system in which $U$ single-antenna UEs transmit data to a BS with $B$ antenna elements. To model this scenario, we use the standard input-output relation presented in (2.1):

$$\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}.$$

Equalization is concerned with forming an estimate $\hat{\mathbf{z}}$ of the transmit signal vector $\mathbf{s}_0$ along with reliability estimates for each entry in $\hat{\mathbf{z}}$. These two quantities are then used to compute hard-output estimates for the transmit symbols or bit-wise soft information in the form of log-likelihood ratios [57, 139]. Consider a general *centralized* equalizer $\{\hat{\mathbf{z}}, \hat{\sigma}^2\} = \mathcal{E}(\mathbf{y}, \mathbf{H})$ that takes the received vector $\mathbf{y}$ and the MIMO channel matrix $\mathbf{H}$ in order to compute (i) an estimate $\hat{\mathbf{z}}$ for the true transmit vector $\mathbf{s}_0$ and (ii) the associated error variance vector $\hat{\sigma}^2$. The error variance vector characterizes the post-equalization residual interference and noise variance on each entry of the estimate $\hat{\mathbf{z}}$. Mathematically, this quantity corresponds to the variances of each entry in the residual interference and noise vector defined as $\mathbf{e} = \hat{\mathbf{z}} - \mathbf{s}_0$, i.e., $\hat{\sigma}^2 = \mathbb{E}_{\mathbf{e}}[\mathbf{e}\mathbf{e}^{\mathrm{H}}]$.

The literature describes a range of linear and nonlinear equalization algorithms for small-scale and massive MU-MIMO data detection [8, 22, 39, 133]. Linear methods, such as MF, ZF, and L-MMSE are among the most common algorithms, mainly due to their simplicity and low computational complexity [22]. Nevertheless, nonlinear equalizers, such as the LAMA algorithm in Algorithm 2, have been shown to (often significantly) outperform linear equalizers at the cost

of higher computational complexity [11,39,153].

### 5.2.2 Basics of Decentralized Equalization

As in [17], we partition the $B$ BS antenna elements into $C \in \{1, 2, \ldots, B\}$ independent *antenna clusters*. The $c$th antenna cluster is associated with $B_c = w_c B$ BS antennas so that $w_c \in [0, 1]$ and $\sum_{c=1}^{C} w_c = 1$. Each cluster contains local RF components and only requires access to local channel state information (CSI) acquired in a local channel estimation (CHEST) unit. Without loss of generality, we partition the receive vector $\mathbf{y} = [\mathbf{y}_1^{\mathrm{T}}, \ldots, \mathbf{y}_C^{\mathrm{T}}]^{\mathrm{T}}$, the channel matrix $\mathbf{H} = [\mathbf{H}_1^{\mathrm{T}}, \ldots, \mathbf{H}_C^{\mathrm{T}}]^{\mathrm{T}}$, and the noise vector $\mathbf{n} = [\mathbf{n}_1^{\mathrm{T}}, \ldots, \mathbf{n}_C^{\mathrm{T}}]^{\mathrm{T}}$ in (2.1). For this antenna partitioning scheme, the input-output relation corresponding to the local receive vector $\mathbf{y}_c$ associated with the $c$th cluster can be written as

$$\mathbf{y}_c = \mathbf{H}_c \mathbf{s}_0 + \mathbf{n}_c, \quad c = 1, \ldots, C, \tag{5.1}$$

with $\mathbf{y}_c \in \mathbb{C}^{B_c}$, $\mathbf{H}_c \in \mathbb{C}^{B_c \times U}$, and $\mathbf{n}_c \in \mathbb{C}^{B_c}$. The following subsections describe two decentralized equalization architectures that compute estimates for the transmit vector $\mathbf{s}_0$ by performing local computations in each antenna cluster using only information of the local receive vector $\mathbf{y}_c$ and channel matrix $\mathbf{H}_c$ followed by fusion of the results from all clusters.

### 5.2.3 Partially Decentralized (PD) Equalization Architecture

The partially decentralized (PD) equalization architecture is illustrated in Fig. 5.1(a). First, each cluster $c = 1, \ldots, C$ independently (and in parallel) preprocesses the local receive vector $\mathbf{y}_c$ and channel matrix $\mathbf{H}_c$ by computing the

$U$-dimensional local MF vector $\mathbf{y}_c^{\text{MF}} = \mathbf{H}_c^{\text{H}}\mathbf{y}_c$ and the $U \times U$ local Gram matrix $\mathbf{G}_c = \mathbf{H}_c^{\text{H}}\mathbf{H}_c$. Second, a *feedforward* adder tree, indicated with the symbol $\oplus$ in Fig. 5.1(a), is used to compute the complete MF vector and Gram matrix as follows:

$$\mathbf{y}^{\text{MF}} = \sum_{c=1}^{C} \mathbf{y}_c^{\text{MF}} \quad \text{and} \quad \mathbf{G} = \sum_{c=1}^{C} \mathbf{G}_c. \tag{5.2}$$

Third, we perform linear or nonlinear equalization in a centralized unit that computes the estimate $\hat{\mathbf{z}} \in \mathbb{C}^U$ and the post-equalization error variance vector $\hat{\sigma}^2 \in \mathbb{C}^U$. The tuple $\{\hat{\mathbf{z}}, \hat{\sigma}^2\}$ is then used to compute hard- or soft-output estimates.

In Section 5.3, we will detail MF, ZF, L-MMSE equalization, and a new LAMA-based equalization algorithm [39] for the PD architecture, all of which directly operate on the $u$-dimensional fused MF vector $\mathbf{y}^{\text{MF}}$ and Gram matrix $\mathbf{G}$. Since the MF vector is a sufficient statistic for the transmit signal $\mathbf{s}_0$ [57], we will show that the PD equalization does not incur a SINR performance loss compared to centralized MF, ZF, L-MMSE, and LAMA equalizers.

### 5.2.4 Fully Decentralized (FD) Equalization Architecture

The PD architecture requires a summation of both the local MF vectors and the local Gram matrices, which involves potentially large amounts of data to be transmitted to the central equalization unit, especially in channels with short coherence time. The fully decentralized (FD) equalization architecture illustrated in Fig. 5.1(b) avoids the transmission of the local Gram matrices altogether at the cost of a (typically small) performance loss. First, each cluster $c = 1, \ldots, C$ independently (and in parallel) performs CHEST, preprocessing, *and* equalization, i.e.,

directly forms a local estimate $\hat{\mathbf{z}}_c \in \mathbb{C}^U$ and local post-equalization error variance vector $\hat{\boldsymbol{\sigma}}_c^2 \in \mathbb{C}^U$. Second, a feedforward fusion tree in Fig. 5.1(b), optimally combines the local estimates $\hat{\mathbf{z}}_c$ using information from the error variance vectors $\hat{\boldsymbol{\sigma}}_c^2$ in order to generate the final output tuple $\{\hat{\mathbf{z}}, \hat{\sigma}^2\}$.

In Section 5.4, we will detail the optimal fusion rule as well as MF, ZF, L-MMSE, and LAMA-based equalization for the FD architecture. We will also provide an SINR performance analysis in the large-system limit.

### 5.2.5 Signal-to-Interference-and-Noise-Ratio (SINR) Analysis

To analyze the performance of linear and nonlinear equalization algorithms for the PD and FD architectures, we will focus on the large-system limit and Rayleigh-fading channels.

**Definition 13** (Rayleigh fading)**.** *A MIMO channel is* Rayleigh fading *if the channel matrix* $\mathbf{H}$ *has i.i.d. circularly symmetric complex Gaussian entries with variance* $1/B$ *per entry.*

By considering the large-system limit and Rayleigh-fading channels, Tse and Hanly have shown in [13] that linear equalizers, such as MF, ZF, and L-MMSE, *decouple* MIMO systems into parallel and independent additive white Gaussian noise (AWGN) channels. This means that the estimate $\hat{\mathbf{z}}$ of such linear equalizers can be modeled on a per-UE basis in a statistically equivalent manner as follows:

$$\hat{z}_u = s_{0,u} + e_u, \quad u = 1, \ldots, U, \tag{5.3}$$

where $e_u \in \mathbb{C}$ represents residual interference and noise. Furthermore, the quantity $e_u$ turns out to be (i) statistically independent of $s_{0,u}$ and (ii) circularly

symmetric complex Gaussian with *decoupled noise variance* $\sigma^2$, which does not depend on the UE index $u$. This result also implies that all entries of the error variance vector $\hat{\boldsymbol{\sigma}}^2$ correspond to $\sigma^2$. In Section 5.3 and Section 5.4 for the PD and FD architecture, respectively, we will build upon this asymptotic analysis framework in order to theoretically characterize the per-UE post-equalization SINR of the decoupled system (5.3):

$$\mathsf{SINR} \triangleq \frac{E_s}{\sigma^2}. \tag{5.4}$$

Numerical results that validate our asymptotic analysis in finite-dimensional systems will be presented in Section 5.5.

## 5.3 Partially Decentralized (PD) Equalization

We start by reviewing linear equalization algorithms for the PD architecture depicted in Fig. 5.1(a), and adapt the well-known Tse-Hanly equations [13] to analyze the associated post-equalization SINR performance in the large-system limit. We then present a new, nonlinear equalization algorithm that builds upon LAMA proposed in [39], and we develop a corresponding SINR performance analysis for the PD architecture.

### 5.3.1 Linear Equalization Algorithms for the PD Architecture

Since the MF output $\mathbf{y}^{\mathrm{MF}}$ in (5.2) is a sufficient statistic for the transmit signal vector $\mathbf{s}_0$, a variety of optimal and suboptimal equalization-based data detection algorithms can be derived from this quantity [57]. For MF-based equalization in

the PD architecture, we use (5.2) to form the estimate

$$\hat{\mathbf{z}}^{\text{MF}} = \text{diag}(\mathbf{G})^{-1}\mathbf{y}^{\text{MF}},$$

where the diagonal matrix $\text{diag}(\mathbf{G})^{-1}$ is computed in the centralized equalization unit; see Fig. 5.1(a). The MF estimate $\hat{\mathbf{z}}^{\text{MF}}$ can then be used to perform either hard- or soft-output data detection. For soft-output data detection, one requires the error variance vector given by

$$\begin{aligned}\hat{\sigma}^2_{\text{MF}} = \text{diag}\big(&\text{diag}(\mathbf{G})^{-1}\mathbf{G}\text{diag}(\mathbf{G})^{-\text{H}}N_0 \\ &+ \Big(\text{diag}(\mathbf{G})^{-1}\mathbf{G} - \mathbf{I}_U\Big)\Big(\text{diag}(\mathbf{G})^{-1}\mathbf{G} - \mathbf{I}_U\Big)^{\text{H}}E_s\big)\end{aligned}$$

that contains the post-equalization SINR for each entry of $\hat{\mathbf{z}}^{\text{MF}}$. Note that MF-based equalization was shown to be optimal (i) for a fixed number of UEs and infinitely many BS antennas [53], which is equivalent to $\beta \to 0$ in the large-system limit, or (ii) in the low-SNR regime [13]. The estimate of the ZF equalizer for the PD architecture is given by

$$\hat{\mathbf{z}}^{\text{ZF}} = \mathbf{G}^{-1}\mathbf{y}^{\text{MF}},$$

where the matrix $\mathbf{G}^{-1}$ is computed in the centralized equalization unit. The associated error variance vector is given by

$$\hat{\sigma}^2_{\text{ZF}} = \text{diag}(\mathbf{G}^{-1})N_0.$$

For L-MMSE equalization, we have

$$\hat{\mathbf{z}}^{\text{L-MMSE}} = (\mathbf{G} + \rho\mathbf{I}_U)^{-1}\mathbf{y}^{\text{MF}}.$$

where the matrix $(\mathbf{G} + \rho\mathbf{I}_U)^{-1}$ is computed in the centralized equalization unit. The L-MMSE regularization parameter is set to $\rho = N_0/E_s$ for complex-valued

constellations (e.g., QPSK or 16-QAM). The associated error variance vector is given by

$$\begin{aligned}\hat{\boldsymbol{\sigma}}^2_{\text{L-MMSE}} = \operatorname{diag}\big((\mathbf{G}+\rho\mathbf{I}_U)^{-1}\mathbf{G}(\mathbf{G}+\rho\mathbf{I}_U)^{-\mathrm{H}}N_0 \\ + \Big((\mathbf{G}+\rho\mathbf{I}_U)^{-1}\mathbf{G}-\mathbf{I}_U\Big)\Big((\mathbf{G}+\rho\mathbf{I}_U)^{-1}\mathbf{G}-\mathbf{I}_U\Big)^{\mathrm{H}}E_s\big).\end{aligned}$$

We reiterate that the MF, ZF, and L-MMSE equalizers for the PD architecture deliver exactly the *same estimates* as their centralized counterparts—the only difference is the way the involved quantities are computed.

As shown in [13] and outlined in Section 5.2.5, centralized MF, ZF, and L-MMSE equalizers decouple MIMO systems in the large-system limit and for Rayleigh fading channels; this implies that the entries of the error variance vectors $\hat{\boldsymbol{\sigma}}^2_{\text{MF}}$, $\hat{\boldsymbol{\sigma}}^2_{\text{ZF}}$, and $\hat{\boldsymbol{\sigma}}^2_{\text{L-MMSE}}$ converge to the decoupled noise variance $\sigma^2$ of the MF, ZF, and L-MMSE equalizer, respectively. Since linear equalizers in the PD architecture yield exactly the same estimates as in a centralized architecture, we can directly characterize the associated decoupled noise variance $\sigma^2_{\text{PD}}$ in the PD architecture using the following result.

**Theorem 26** ([13, Thm. 3.1])**.** *Fix the system ratio $\beta = U/B$, and assume the large-system limit and Rayleigh fading channels. Then, the decoupled noise variance $\sigma^2_{PD}$ for MF, ZF, and L-MMSE equalization in a centralized or PD architecture, is the solution to the following fixed-point equation*

$$\sigma^2_{PD} = N_0 + \beta\Psi(\sigma^2_{PD}), \tag{5.5}$$

*where the MSE function $\Psi(\sigma^2)$ is given by*

$$\Psi(\sigma^2) = E_s, \tag{MF}$$

$$\Psi(\sigma^2) = \sigma^2, \tag{ZF}$$

$$\Psi(\sigma^2) = \frac{E_s}{E_s+\sigma^2}\sigma^2, \tag{L-MMSE}$$

*for MF, ZF, and L-MMSE equalization, respectively.*

We note that the expression for the ZF equalizer only holds for $\beta < 1$, whereas the expressions for MF and L-MMSE hold for general system ratios $\beta$.[1] From Theorem 26, we obtain closed-form expressions for the post-equalization SINR in (5.4) for MF, ZF, and L-MMSE equalization in the PD architecture.

**Corollary 27.** *Assume that the conditions of Theorem 26 hold. Then, the post-equalization* SINR *for MF, ZF, and L-MMSE equalization in the PD architecture are given by*

$$\mathsf{SINR}_{\mathrm{PD}}^{\mathrm{MF}} = \frac{E_s/N_0}{1+\beta E_s/N_0}, \tag{MF}$$

$$\mathsf{SINR}_{\mathrm{PD}}^{\mathrm{ZF}} = \frac{E_s}{N_0}(1-\beta), \text{ for } \beta < 1, \tag{ZF}$$

$$\mathsf{SINR}_{\mathrm{PD}}^{\text{L-MMSE}} = \frac{1}{2}\left(\sqrt{\left(1-\frac{E_s}{N_0}(1-\beta)\right)^2 + 4\frac{E_s}{N_0}} - \left(1-\frac{E_s}{N_0}(1-\beta)\right)\right). \tag{L-MMSE}$$

We note that in the massive MU-MIMO regime, which corresponds to $\beta \to 0$, all post-equalization SINR expressions converge to $E_s/N_0$, which confirms the well-known fact that MF is optimal in this scenario [53]. It can also be shown that $\mathsf{SINR}_{\mathrm{PD}}^{\text{L-MMSE}}$ bounds $\mathsf{SINR}_{\mathrm{PD}}^{\mathrm{MF}}$ and $\mathsf{SINR}_{\mathrm{PD}}^{\mathrm{ZF}}$ from above for all system ratios and in all SNR regimes. Hence, L-MMSE equalization is often the preferred choice in realistic massive MU-MIMO systems [6,22]. We reiterate that the SINR expressions listed in Corollary 27 are also valid for centralized architectures.

[1] The asymptotic SINR performance of ZF equalization via the Moore-Penrose pseudo inverse when $\beta \geq 1$ was analyzed in [137].

### 5.3.2 LAMA for the PD Architecture

The LAMA algorithm presented in Algorithm 2 is a nonlinear equalizer is able to achieve individually-optimal performance in the large-system limit given certain conditions on the antenna ratio $\beta$ and the noise variance $N_0$ are satisfied. LAMA operates directly on the input-output relation in (2.1) and is, hence, designed for centralized processing. We now develop a novel variant of LAMA that directly operates on the complete MF output $\mathbf{y}^{\mathrm{MF}}$ and the Gram matrix $\mathbf{G}$ in (5.2) to enables its use in the PD architecture. Since the antenna configuration in massive MU-MIMO systems typically satisfies $U \ll B$, the LAMA-PD algorithm operates on a lower dimension which reduces complexity while delivering exactly the same estimates as the original LAMA algorithm. We note that LAMA was derived in the large-system limit and for Rayleigh fading channels [87], but these assumptions are not required in practice. We next summarize the LAMA-PD algorithm; the derivation can be found in Appendix A.3.1.

**Algorithm 4** (LAMA-PD)**.** *Initialize* $s_\ell = \mathbb{E}_S[S]$ *for* $\ell = 1, \ldots, U$, $\phi^{(1)} = \mathbb{V}\mathrm{ar}_S[S]$, *and* $\mathbf{v}^{(1)} = \mathbf{0}_{B\times 0}$. *Then, for every iteration* $t = 1, 2, \ldots, T_{max}$, *compute the following steps:*

$$\begin{aligned} \hat{\mathbf{z}}^t &= \mathbf{y}^{\mathrm{MF}} + (\mathbf{I} - \mathbf{G})\hat{\mathbf{s}}^t + \mathbf{v}^t \qquad (5.6)\\ \hat{\mathbf{s}}^{t+1} &= \mathsf{F}(\hat{\mathbf{z}}^t, N_0 + \hat{\tau}_t^2) \\ \hat{\tau}_{t+1}^2 &= \beta\langle \mathsf{G}(\hat{\mathbf{z}}^t, N_0 + \hat{\tau}_t^2)\rangle \\ \mathbf{v}^{t+1} &= \frac{\hat{\tau}_{t+1}^2}{N_0 + \hat{\tau}_t^2}(\hat{\mathbf{z}}^t - \mathbf{s}^t). \end{aligned}$$

*The estimates and error variances of LAMA are* $\hat{\mathbf{z}}^t$ *and* $\sigma_{t,\mathrm{LAMA}}^2 = N_0 + \hat{\tau}_t^2$, *respectively.*

In order to analyze the post-equalization SINR of LAMA-PD, it is key to real-

ize that the equalization output $\hat{\mathbf{z}}^t$ is equivalent to that of the original centralized LAMA algorithm in Algorithm 2. As discussed in Section 3.3.3, LAMA (and hence LAMA-PD) decouples the MIMO system into parallel AWGN channels.

For $t \to \infty$, the SE recursion in Theorem 2 converges to the same fixed-point equation of linear equalizers in (5.5), where the only difference is the MSE function (3.13). As for linear equalizers, we can use the fixed-point equation in (5.5). Correspondingly, we can use SE to analyze the post-equalization SINR performance of LAMA and LAMA-PD. Unfortunately, there are no closed-form expressions known for the decoupled noise variance or the SINR for LAMA and LAMA-PD with discrete constellations, due to the specific form of the MSE function (3.13). Nevertheless, we can numerically compute (3.13) and, hence, analyze the SINR. A corresponding SINR comparison with linear equalizers is given in Section 5.5.

## 5.4 Fully Decentralized (FD) Equalization

We next discuss optimal fusion for linear and nonlinear equalization in the PD architecture as depicted in Fig. 5.1(b). We then analyze the post-equalization SINR of the proposed equalizers depending on the antenna cluster allocation strategy.

### 5.4.1 Optimal Fusion for the FD Architecture

As detailed in Section 5.2.4, each cluster $c = 1, \ldots, C$ in the FD architecture independently computes a local estimate $\hat{\mathbf{z}}_c$ and associated error variance vec-

tor $\hat{\boldsymbol{\sigma}}_c^2$. Then, the vectors $\hat{\mathbf{z}}_c$ and $\hat{\boldsymbol{\sigma}}_c^2$ for $c = 1, \ldots, C$ are fused to compute the final output tuple $\{\hat{\mathbf{z}}, \hat{\boldsymbol{\sigma}}^2\}$. Since in the large-system limit and for Rayleigh fading channels, the considered equalizers decouple MIMO systems into parallel and independent AWGN channels (see Section 5.2.5), we focus on *linear* fusion of the local estimates in Fig. 5.1(b). Specifically, the proposed FD architecture computes the fused estimate $\hat{\mathbf{z}}$ by combining the local estimates for each UE as follows:

$$\hat{z}_u = \sum_{c=1}^{C} \nu_{c,u} \hat{z}_{c,u}, \quad u = 1, \ldots, U. \tag{5.7}$$

Here, $\hat{z}_{c,u}$ is the local estimate for $u$th UE at cluster $c$ and the weights $\nu_{c,u}$, $u = 1, \ldots, U$, depend on the per-cluster error variance vector $\hat{\boldsymbol{\sigma}}_c^2$. In what follows, we are interested in the optimal set of weights for the following criterion.

**Definition 14** (Optimal fusion)**.** *Optimal fusion for the FD architecture maximizes the per-UE post-equalization SINR of the final estimate $\hat{z}_u$ obtained from (5.7) while $\sum_{c=1}^{C} \nu_{c,u} = 1$.*

In other words, optimal fusion defines a set of weights $\nu_{c,u}$, $c = 1, \ldots, C$, $u = 1, \ldots, U$, so that the decoupled noise variances contained in $\hat{\boldsymbol{\sigma}}^2$ associated with the fused estimate $\hat{\mathbf{z}}$ are minimized. The following result summarizes the optimal fusion rule; a short proof is given in Appendix A.3.2.

**Lemma 28.** *Let $\sigma_{c,u}^2$, $c = 1, \ldots, C$, $u = 1, \ldots, U$, be a set of given error variances for UE u and cluster c. Assume that the residual interference and noise terms are zero mean and uncorrelated among the clusters. Then, the weights that yield optimal fusion according to Definition 14 are given by*

$$\nu_{c,u} = \frac{1}{\sigma_{c,u}^2} \left( \sum_{c'=1}^{C} \frac{1}{\sigma_{c',u}^2} \right)^{-1}, \tag{5.8}$$

*for $c = 1, \ldots, C$ and $u = 1, \ldots, U$.*

### 5.4.2 SINR Analysis of Optimal Fusion in the FD Architecture

We are now interested in analyzing the post-fusion SINR for the FD architecture in the large-system limit. The following theorem analyzes the decoupled noise variance $\sigma_c^2$ for each cluster $c = 1, \ldots, C$; the proof is given in Appendix A.3.3.

**Lemma 29.** *Assume MF, ZF, L-MMSE, or LAMA equalization in each cluster $c = 1, \ldots, C$. Consider the large-system limit and Rayleigh fading channels. Then, the input-output relation of each cluster is decoupled into parallel channels of the form (5.3) with decoupled noise variance $\sigma_c^2$ given by a solution to the following fixed-point equation:*

$$w_c \sigma_c^2 = N_0 + \beta \Psi(\sigma_c^2).$$

*Here, $\Psi(\sigma_c^2)$ is the MSE function of the equalizer in cluster $c$.*

This result shows that the per-UE error variances $\sigma_{c,u}^2$ will become independent of the UE index $u$ in the large-system limit and for Rayleigh-fading channels. Furthermore, the decoupled noise variances $\sigma_c^2$ depend on the fraction $w_c$ of BS antennas associated with cluster $c$.

The following result establishes the post-fusion SINR in the FD architecture; a short proof is given in Appendix A.3.4.

**Theorem 30.** *Let the assumptions of Lemma 29 hold and $\sigma_c^2$, $c = 1, \ldots, C$, be the per-cluster decoupled noise variances. Then, the decoupled noise variance $\sigma_{FD}^2$ of the fused estimate in (5.7) of the FD architecture is given by*

$$\sigma_{FD}^2 = \left( \sum_{c=1}^{C} \frac{1}{\sigma_c^2} \right)^{-1} = N_0 + \beta \sum_{c=1}^{C} \nu_c \Psi(\sigma_c^2). \tag{5.9}$$

We note that this result implies that the post-fusion SINR with optimal fusion according to Definition 14, denoted by $\mathsf{SINR}_{\mathrm{FD}}$, corresponds to the sum of the per-cluster SINR values.

Finally, we have the following intuitive result which implies that for a given equalizer, the FD architecture cannot outperform the PD architecture; the proof is given in Appendix A.3.5.

**Lemma 31.** *Let $N_0 > 0$ and assume the large-system limit and Rayleigh-fading channels. Then, the output SINR for the FD and PD architectures satisfy $\mathsf{SINR}_{\mathrm{FD}} \leq \mathsf{SINR}_{\mathrm{PD}}$. Equality holds for $\beta \to 0$, $C = 1$, or if MF-based equalization is used.*

### 5.4.3 Antenna Partitioning Strategies for Linear Equalizers

We now analyze the post-fusion SINR performance of linear algorithms for the FD architecture, depending on the antenna allocation strategy, i.e., on the fraction of antennas $w_c$ used per cluster $c$. For the following analysis, we assume the large-system limit and Rayleigh fading channels.

For MF with the FD architecture, the post-fusion SINR is equivalent to that of the PD architecture (and that of centralized processing), as shown in Lemma 31. Hence, the antenna partitioning strategy does not affect the SINR performance.

For ZF equalization in the FD architecture, we have the following result; the proof is given in Appendix A.3.6.

**Lemma 32.** *Assume that the B BS antennas are divided into C clusters so that $w_c \geq \beta$ holds for $c = 1, \ldots, C$. Then, the post-fusion SINR for ZF equalization is given by*

$$\mathsf{SINR}_{\mathrm{FD}}^{\mathrm{ZF}} = \frac{E_s}{N_0}(1 - C\beta). \tag{5.10}$$

Interestingly, we observe that the post-fusion SINR $\mathsf{SINR}_{\mathrm{FD}}^{\mathrm{ZF}}$ does not depend on the antenna allocation strategy; this implies that the per-cluster antenna frac-

tion $w_c$ can be chosen arbitrarily as long as $w_c \geq \beta$ for $c = 1, \ldots, C$.[2] Note, however, that equally-sized clusters are desirable in practice as they may minimize the interconnect or chip I/O bandwidth as well as the computational complexity per computing fabric.

For L-MMSE equalization in the FD architecture, we have the following result; the proof is given in Appendix A.3.7.

**Lemma 33.** *Assume that the B BS antennas are divided into C clusters so that* $\sum_{c=1}^{C} w_c = 1$ *holds with* $w_c \geq 0$ *for* $c = 1, \ldots, C$. *Then, the post-fusion SINR of the L-MMSE equalizer is given by*

$$\begin{aligned} \mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}} = & \frac{1}{2} \sum_{c=1}^{C} \sqrt{\left(1 - \frac{E_s}{N_0}(w_c - \beta)\right)^2 + 4\frac{E_s}{N_0} w_c} \\ & - \frac{1}{2}\left(C - \frac{E_s}{N_0}(1 - C\beta)\right) \end{aligned} \tag{5.11}$$

We see from Lemma 33 that the post-fusion SINR expression for L-MMSE equalization depends on the antenna allocation strategy, i.e., on the weights $w_c$, which is in contrast to ZF equalization (cf. Lemma 32). In addition, L-MMSE equalization does not require the restriction $w_c \geq \beta$ for ZF-equalization as the post-fusion SINR expression holds even for underdetermined systems [13]. Hence, it is natural to ask what the optimal cluster allocation strategy is. The following result is rather disappointing; the proof is given in Appendix A.3.8.

**Lemma 34.** *The cluster allocation strategy that maximizes the post-fusion SINR for the L-MMSE equalizer* $\mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}}$ *is* $w_c = 1$ *for some c and* $w_{c'} = 0$ *for* $c' \neq c$.

Clearly, without any systematic requirements on the cluster ratios $w_c$, $c = 1, \ldots, C$, besides $w_c \geq 0$, maximizing $\mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}}$ so that $\sum_{c=1}^{C} w_c = 1$ corre-

[2]The condition $w_c \geq \beta$ implies that $\sum_{c=1}^{C} w_c = 1 \geq C\beta$ so the SINR expression in Lemma 32 is well-defined.

sponds to a centralized architecture, i.e., all antennas should be allocated to a single cluster. Since the key idea of DBP was to mitigate interconnect and I/O bandwidth as well as computation bottlenecks, such an optimal allocation strategy is undesirable in practice. We next show that the most desirable (from a practical viewpoint) cluster allocation strategy, i.e., one for which all clusters have an equal number of antennas, yields the worst post-fusion SINR; the proof is given in Appendix A.3.9

**Lemma 35.** *Assume that the B BS antennas are divided into C clusters so that* $\sum_{c=1}^{C} w_c = 1$ *with* $w_c \geq 0$, $c = 1, \ldots, C$. *Then, we have the following lower bound on the post-fusion SINR:*

$$\mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}} \geq \frac{1}{2}\sqrt{\left(1 - \frac{E_s}{N_0}(1 - C\beta)\right)^2 + 4\frac{E_s}{N_0}C} - \frac{1}{2}\left(C - \frac{E_s}{N_0}(1 - C\beta)\right). \tag{5.12}$$

*Furthermore, the lower bound is achieved with equality if the antennas are distributed uniformly across all clusters, i.e., where* $w_c = 1/C$ *for* $c = 1, \ldots, C$.

We conclude by noting that even though uniform cluster sizes are the worst for L-MMSE equalization in the FD architecture, L-MMSE equalization outperforms ZF equalization for all possible partitioning schemes in terms of the post-fusion SINR. Hence, L-MMSE equalization is desirable in practice.

## 5.5 Numerical Results

We now investigate the performance of decentralized equalization in the large-system limit and for Rayleigh-fading channels using the SINR expressions from

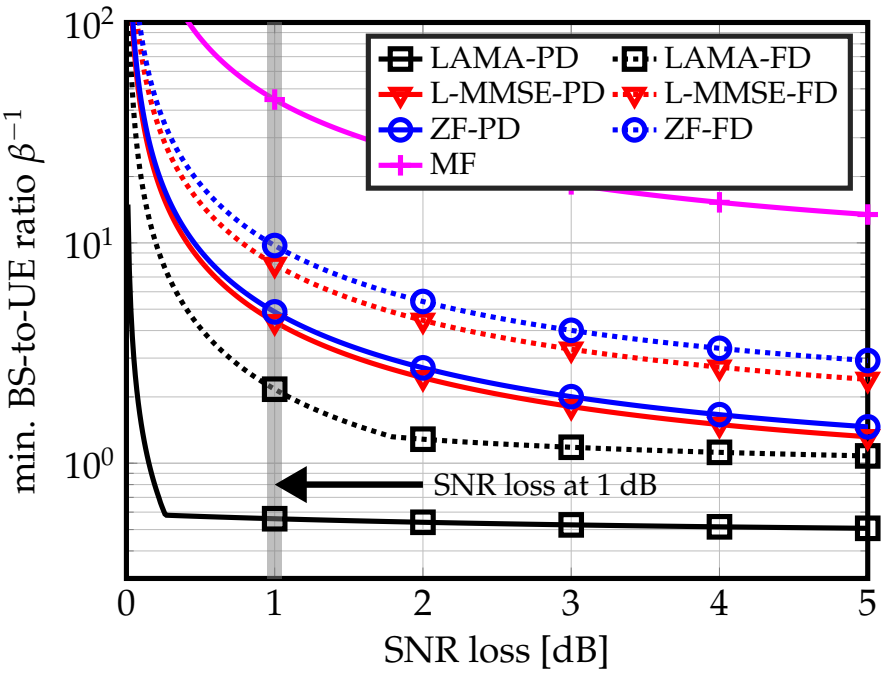



Figure 5.2: Achievable rate analysis of DBP with feedforward architectures for QPSK constellation. We fix the target achievable rate $R = 1.99$ and plot the required minimum BS-to-UE antenna ratio $\beta^{-1}$ versus SNR loss. LAMA outperforms MF, ZF, and L-MMSE in both PD and FD architectures. LAMA-FD outperforms L-MMSE-PD in terms of required minimum BS-to-UE ratio at all SNR loss.

Sections 5.3 and 5.4. We show error-rate simulation results to validate our asymptotic results in finite-dimensional systems. We also provide results for an LTE-like massive MU-MIMO system to demonstrate the efficacy of our solutions in a realistic scenario.

### 5.5.1 Achievable Rate Analysis

We first investigate the achievable rates of our feedforward architectures with focus on the large-system limit and Rayleigh fading channels. We consider $C = 2$ clusters with uniform antenna partitioning, i.e., $w_c = 1/C$ for $c = 1, \ldots, C$. We

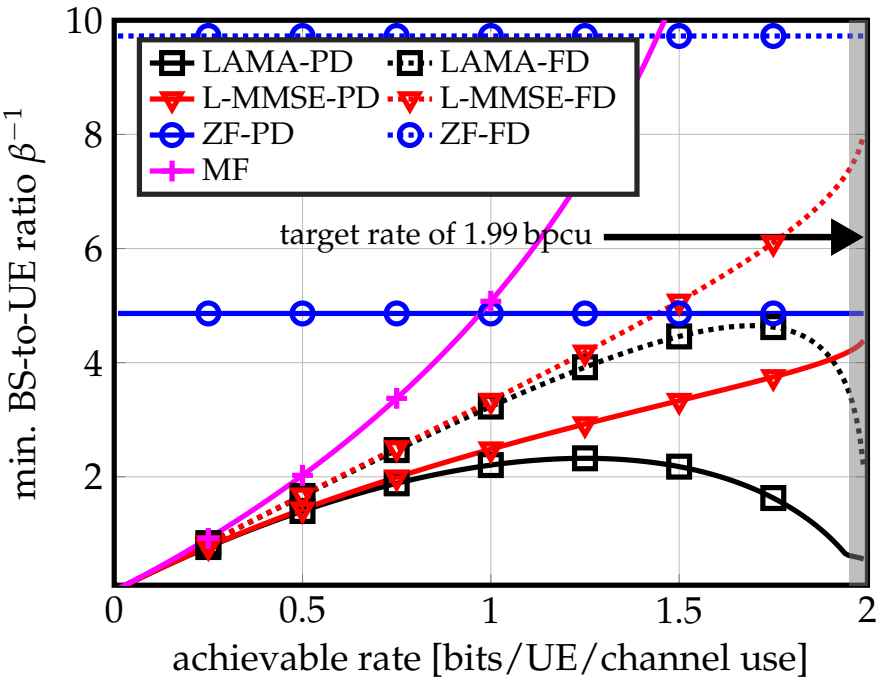



Figure 5.3: Achievable rate analysis of DBP with feedforward architectures for QPSK constellation. We fix the SNR loss to 1 dB and plot the required minimum BS-to-UE antenna ratio $\beta^{-1}$ versus achievable rate. LAMA outperforms MF, ZF, and L-MMSE in both PD and FD architectures.

define the average receive signal-to-noise ratio (SNR) as $\mathsf{SNR} = \beta E_s / N_0$ and use an interference-free AWGN channel with variance $N_0$ as the baseline, which coincides to the large-system limit of massive MU-MIMO systems with $\beta \to 0$. Concretely, we will use the following performance metric.

**Definition 15** (SNR loss)**.** *We define the SNR loss of an equalizer as the excess SNR required to achieve the same target rate $R$ of an interference-free AWGN channel with variance $N_0$.*

In Fig. 5.2, we use QPSK and a target rate of $R = 1.99$ bits/UE/channel use; in Fig. 5.4 we use 16-QAM and a target rate of $R = 3$ bits/UE/channel use. Both figures investigate the minimum required BS-to-UE ratio $\beta^{-1}$ for a given SNR loss, which characterizes how many more BS antennas are required by a

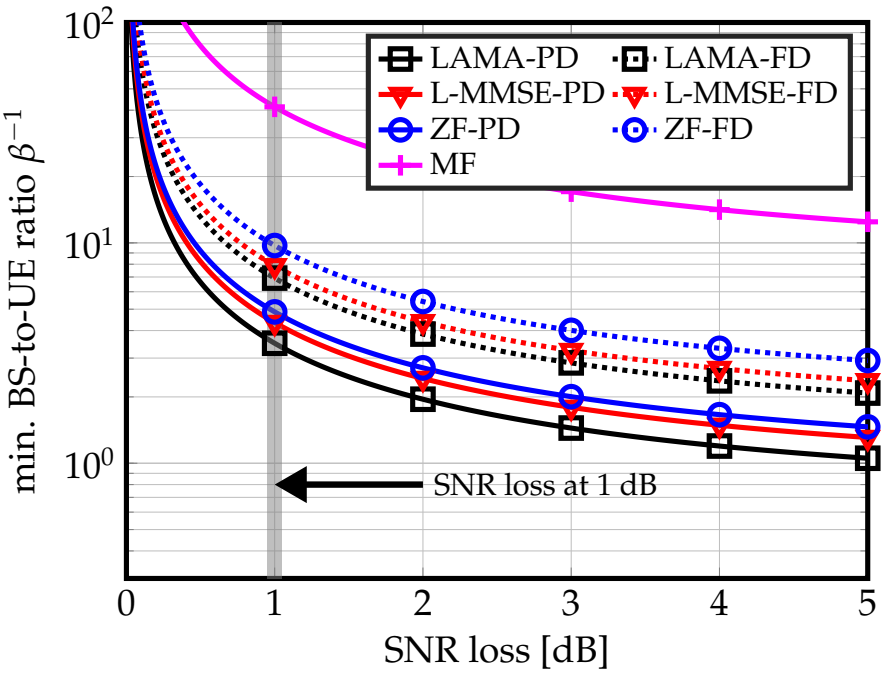



Figure 5.4: Achievable rate analysis of DBP with feedforward architectures for 16-QAM constellation. We fix the target achievable rate $R = 3$ and plot the required minimum BS-to-UE antenna ratio $\beta^{-1}$ versus SNR loss. LAMA outperforms MF, ZF, and L-MMSE in both PD and FD architectures.

given equalizer and feedforward architecture to be able to approach AWGN performance up to a given SNR gap. We observe that for a small SNR loss (i.e., when achieving similar performance as that of an interference-free AWGN channel), we require significantly more BS antennas than UEs, irrespective of the algorithm and architecture. For an SNR loss of 1 dB (shown by a thick vertical line in Figures 5.2 and 5.4), we see that the PD architecture outperforms the FD architecture; this fact is more pronounced in the QPSK scenario as we are trying to achieve 99.5% of the maximum possible rate of 2 bits/UE/channel use for QPSK, whereas for 16-QAM, we are only trying to achieve 75% of the maximum rate. We also see that LAMA-PD significantly outperforms linear equalizers in the PD and FD architectures; MF requires significantly higher BS-to-UE antenna ratios. Interestingly, for QPSK, LAMA-FD significantly outperforms linear equalization

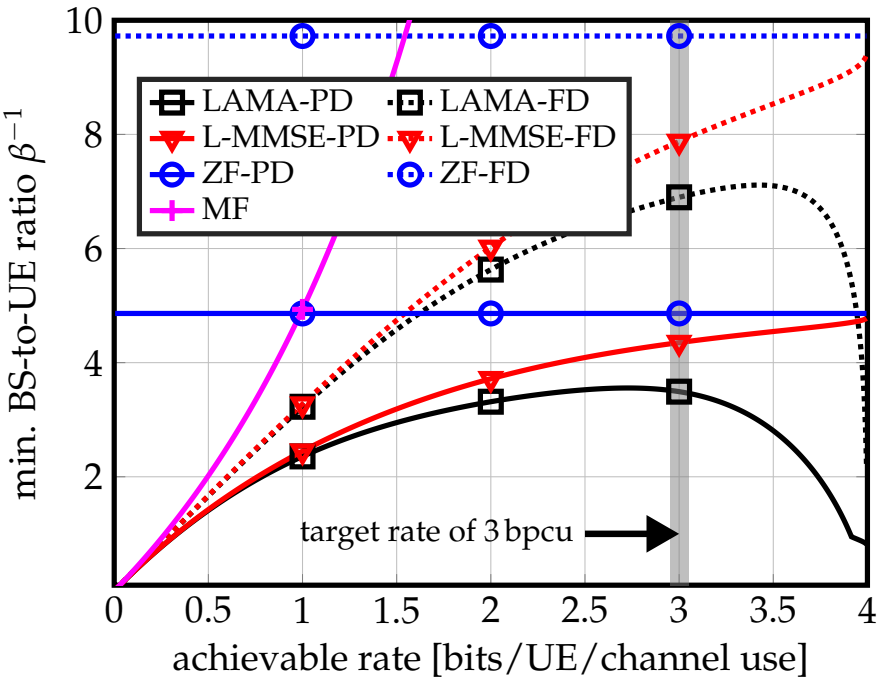



Figure 5.5: Achievable rate analysis of DBP with feedforward architectures for 16-QAM constellation. We fix the SNR loss to 1 dB and plot the required minimum BS-to-UE antenna ratio $\beta^{-1}$ versus achievable rate. LAMA outperforms MF, ZF, and L-MMSE in both PD and FD architectures.

algorithms for the PD architecture in Fig. 5.2 but performs strictly worse for 16-QAM in Fig. 5.4; this is due to the fact that the system-ratio threshold for LAMA to achieve individually-optimal performance is higher for QPSK than for 16-QAM [39]. In summary, LAMA-FD achieves similar performance as linear equalizers with the PD architecture while reducing interconnect and chip I/O bandwidths.

In Fig. 5.3 and Fig. 5.5, we fix the SNR loss to 1 dB and plot the minimum BS-to-UE ratio $\beta^{-1}$ and varying achievable rates for QPSK and 16-QAM, respectively. For both constellations, MF performs equally well than all other methods in the low-rate regime; note that the low-rate regime translates to the low-SNR regime for which MF is known to be optimal. For higher rates, however, MF requires significantly higher BS-to-UE antenna ratios compared to L-MMSE

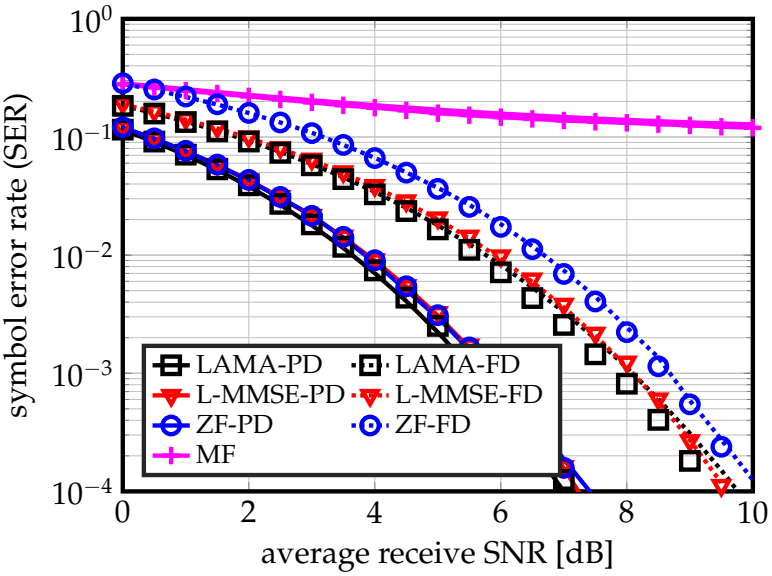



Figure 5.6: Symbol error rate (SER) comparison of equalization in the large-system limit (indicated with lines) and the simulated performance (indicated by the markers) in a $B = 256$ BS antenna $u = 16$ UE massive MU-MIMO system with Rayleigh-fading. Evaluating the SER using our analytical SINR expressions for the large-system limit closely matches numerical simulations in finite-dimensional systems for all considered equalizers and architectures.

or LAMA-based equalization. The PD architecture significantly outperforms the FD architecture for all equalizers, which implies that for high-rates the PD architecture is the preferred choice. Interestingly, the minimum BS-to-UE antenna ratio remains constant for ZF; this implies that as long as one operates below a certain antenna ratio $\beta^*$, ZF is able to support all transmission rates; see [152] for additional details on this behavior. Finally, we see that the minimum BS-to-UE ratio $\beta^{-1}$ decreases for LAMA-FD and LAMA-PD at high rates; this behavior is due to the fact that LAMA in overloaded systems is particularly robust at low and high values of SNR (see [39] for a detailed discussion).

### 5.5.2 Asymptotic vs. Finite-Dimensional Systems

In Fig. 5.6, we compare our analytical SINR expressions in the large-system limit to those in a finite-dimensional massive MU-MIMO scenario. Specifically, we use the SINR from (5.3) in a decoupled AWGN channel to analytically compute the symbol error rate (SER) as well as the simulated SER in an uncoded $B = 256$ BS antenna, $u = 16$ UE massive MU-MIMO system with Rayleigh fading channels. We consider $C = 8$ clusters with uniform antenna partitioning, i.e., $w_c = 1/8$ for all $C = 8$ clusters. First, we observe that the simulated results (indicated with markers) closely match our analytical expressions (indicated with lines). Second, we see that the PD architecture significantly outperforms the FD architecture for $C = 8$ clusters and LAMA outperforms ZF and L-MMSE for both architectures. Third, we see that MF yields poor SER performance, which is due to the fact that MF requires extremely high BS-to-UE antenna ratios to support 4 bits/UE/channel use for 16-QAM; see also Fig. 5.5.

### 5.5.3 Coded Error-rate Performance in Realistic Systems

In Fig. 5.7, we investigate the coded packet error-rate (PER) in a realistic LTE-like massive MU-MIMO system with $B = 64$ BS antennas and $U = 16$ UEs. We consider $C = 2$ clusters with uniform antenna partitioning. We use OFDM with 2048 subcarriers (1200 used for data transmission) with 16-QAM, 14 OFDM symbols per packet, and use a weak rate-5/6 convolutional code with soft-input Viterbi decoding. We consider a WINNER II channel model in an outdoor-to-indoor scenario. For LAMA-PD and FD, we use 10 iterations and perform message damping to mitigate performance loss for finite-dimensional systems

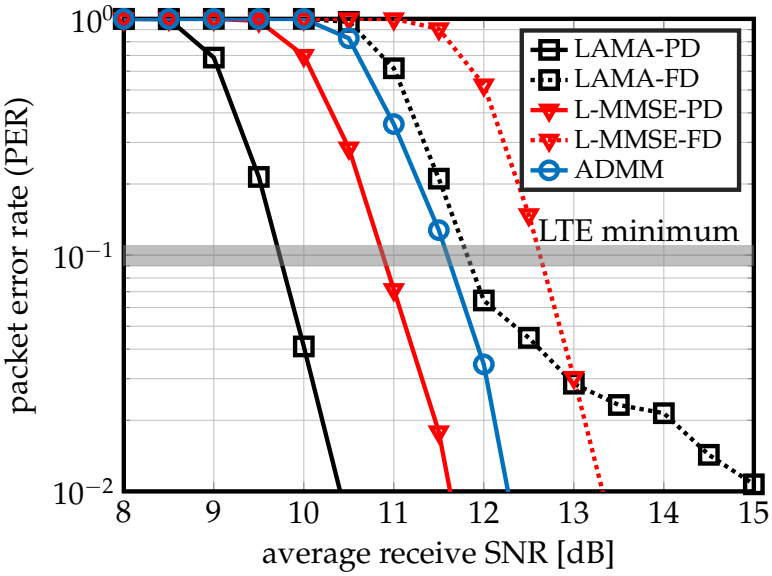



Figure 5.7: Packet error-rate (PER) of an LTE-like massive MU-MIMO-OFDM system with $B = 64$ and $U = 16$. LAMA for the PD and FD architectures clearly outperforms L-MMSE while meeting the 10% LTE minimum PER requirement. LAMA-PD or L-MMSE clearly outperform the consensus-based ADMM method from [17] (which suffers from transfer latency), whereas LAMA-FD closely approaches the performance at minimal lower latency overheads.

with correlated MIMO channel matrices [157]. We also compare LAMA and L-MMSE to the consensus-based ADMM method for DBP proposed in [17], where we use 10 iterations. First, we see that LAMA-PD outperforms all other equalization algorithms by a significant margin, when considering the LTE minimum PER specification of 10%. Second, we observe that the consensus-sharing ADMM method performs slightly better than that of LAMA-FD. The ADMM-based method, however, requires iterative consensus exchange among the clusters which results in low throughput.

## 5.6 Conclusions

We have presented two feedforward architectures for decentralized equalization in massive MU-MIMO systems that mitigate the interconnect and I/O bandwidth bottlenecks and enable parallel processing on multiple computing fabrics. For the two proposed architectures, we have presented linear and nonlinear equalization algorithms, and we have analyzed their post-equalization SINR performance in the large-system limit. We have also performed numerical simulations that confirm our analysis. Our results indicate that nonlinear equalizers are able to achieve near-optimal SINR performance while enabling decentralized computations and low communication overhead among the antenna clusters. Linear equalizers perform equally well for scenarios in which the number of BS antennas is significantly larger than the number of UEs or for systems that use strong coding or low data rates.

CHAPTER 6

# EFFICIENT PREPROCESSING FOR MASSIVE MU-MIMO

## 6.1 Introduction

Massive multi-user (MU) multiple-input multiple-output (MIMO) will be a key technology in fifth-generation (5G) wireless systems [5,53]. The idea of massive MU-MIMO is to equip the base-station (BS) with hundreds of antenna elements while serving tens of user equipments (UEs) in the same time-frequency resource. Such large antenna arrays enable extremely fine-grained beamforming in the uplink (UEs transmit to the BS) and in the downlink (BS transmits to the UEs), which offers superior spectral efficiency compared to traditional, small-scale MIMO technology that use only a few antennas at the BS.

In the uplink, linear data-detection algorithms that rely on minimum-mean square error (MMSE) equalization or zero-forcing (ZF) equalization are known to achieve near-optimal error-rate performance in realistic massive MU-MIMO systems with a finite number of transmit antennas [5,6,57]. Non-linear data-detection algorithms [39,111,153] have recently been shown to outperform linear methods in systems where number of UEs is comparable to the number of BS antennas. Most of these linear and non-linear data-detection algorithms entail high computational complexity, often dominated by the computation of the so-called *Gram matrix* $\mathbf{G} = \mathbf{H}^{\mathrm{H}}\mathbf{H}$ [22,23]. Here, $\mathbf{H} \in \mathbb{C}^{B \times U}$ is the (uplink) channel matrix, $B$ is the number of BS antennas, and $U$ is the number of (single-antenna) users. The computational complexity is orders-of-magnitude higher in wideband systems that use orthogonal frequency-division multiplexing (OFDM) or single-carrier frequency-division multiple access (SC-FDMA), in which a Gram matrix must

be computed for each active subcarrier (i.e., subcarriers used for pilots or data transmission) [22]. For example, Gram matrix computation requires more than $2\times$ higher complexity than data detection for a 128 BS antenna 16 UE antenna MU-MIMO system [23, Table 4.2], and is much higher for a system with more BS antennas. In the massive MU-MIMO downlink, precoding is necessary to focus the transmit energy towards the UEs and to mitigate multi-user interference [5]. In wideband systems, the complexity of linear precoding algorithms is—analogously to the uplink—dominated by Gram matrix computation on all active subcarriers.

While some data-detection and precoding algorithms have been proposed that avoid the computation of the Gram matrix altogether (see, e.g., [8,31,158]), these methods do not allow the re-use of intermediate results in time-division duplexing (TDD) systems. Specifically, the Gram matrix and its inverse cannot be re-used in the uplink (for equalization) and downlink (for precoding), which would significantly lower the computational complexity. Hence, such algorithms inevitably perform redundant computations during data-detection and precoding, which leads to inefficient transceiver designs.

### 6.1.1 Interpolation-Based Matrix Computations

In practical wideband communication systems, e.g., building upon IEEE 802.11n [47] and 3GPP-LTE [159], the channel's delay spread is often substantially smaller than the number of active subcarriers. Hence, the channel coefficients are correlated across subcarriers. This property can be exploited to reduce the computational complexity of commonly-used matrix computations required in

multi-antenna systems. More specifically, the papers [24–27] avoid a brute-force approach in traditional, small-scale, and point-to-point MIMO-OFDM systems by using exact interpolation-based algorithms for matrix inversion and QR factorization. While a few hardware designs [160, 161] have demonstrated the efficacy of these exact interpolation methods in small-scale MIMO systems, their complexity does not scale well to wideband massive MU-MIMO systems with hundreds of BS antennas, tens of users, and thousands of subcarriers; in fact, recent 3GPP specifications on New Radio (NR) access technology shows that the number of active subcarriers is 3300 or 6600 in Rel-15 [162, 163]. In addition, the impact of imperfect channel-state information (CSI) and antenna correlation on such exact, interpolation-based matrix computation algorithms is routinely ignored, but significantly affects their performance in practical scenarios (see Section 6.6 for a detailed discussion).

### 6.1.2 Relevant Prior Art

Data detection and precoding for small-scale, single- and multi-carrier MIMO systems is a well studied topic; see e.g. [57, 139, 164–166] and the references therein. Data detection and precoding algorithms for massive MIMO systems have been proposed in, e.g., [22, 28, 29, 39, 111, 167], which leverage the fact that the Gram matrix is diagonally dominant [64]. However, all of these results (i) do not exploit specific properties of massive MU-MIMO systems and (ii) ignore the fact that time-division duplexing (TDD)-based systems must perform data-detection *and* precoding, and hence, can re-use intermediate results (such as the Gram matrix) to reduce the computational complexity. In contrast, our results exploit the specifics of massive MU-MIMO systems, namely channel hardening,

and enable a re-use of the computations carried out in the uplink for downlink precoding.

The recent report [168] proposed an approximate interpolation-based ZF-based equalizer for a wideband massive MU-MIMO testbed. In contrast to our work, the authors interpolate the *inverse* of the Gram matrix. While simulation results in [168] show that the method works well in practice, no theoretical results have been provided. In contrast, we use approximate methods to interpolate the Gram matrix, and we provide exact analytical results that provide a solid foundation of approximate interpolation methods in massive MU-MIMO systems.

## 6.2 Prerequisites

We start by summarizing the considered wideband massive MU-MIMO system and channel model. We then outline computationally-efficient ways for linear data detection and precoding that make use of the Gram matrix.

### 6.2.1 System Model

Without loss of generality, we focus on the uplink[1] of a wideband massive MU-MIMO system with $B$ base-station antennas, $U$ single-antenna UEs (with $U \ll B$), and $W$ subcarriers. For each active subcarrier $\omega \in \Omega$ with $\Omega$ containing

[1] By assuming channel reciprocity [5], our results directly apply to the downlink, in which $\mathbf{H}^{\mathrm{T}}$ is the downlink channel matrix. Once all of the Gram matrices have been computed, they can be re-used in the downlink for linear (e.g., zero-forcing or Wiener filter) precoding. See Section 6.2.3 for the details.

the indices of the active (data and pilot) subcarriers, we model the received frequency-domain (FD) signal as follows:

$$\mathbf{y}_\omega = \mathbf{H}_\omega \mathbf{s}_\omega + \mathbf{n}_\omega. \tag{6.1}$$

Here, $\mathbf{H}_\omega \in \mathbb{C}^{B \times U}$ is the FD channel matrix, $\mathbf{s}_\omega \in \mathbb{C}^U$ is the transmit vector, and $\mathbf{n}_\omega \in \mathbb{C}^B$ models additive noise. The FD input-output relation in (6.1) is able to model both OFDM and SC-FDMA systems. For OFDM systems, the entries of the transmit vector $\mathbf{s}_\omega$ are taken from a discrete constellation set $\mathcal{O}$ (e.g., 16-QAM); in SC-FDMA systems, the constellation points are assigned in the time-domain and the resulting vectors are transformed into the FD to obtain the transmit vectors $\mathbf{s}_\omega$. See, e.g., [22], for details on SC-FDMA transmission.

### 6.2.2 Wideband Channel Model

In wideband MIMO multicarrier systems, the FD channel matrices $\mathbf{H}_\omega$, $\omega = 0, \ldots, W-1$, are directly related to the time-domain (TD) matrices $\widehat{\mathbf{H}}_\ell \in \mathbb{C}^{B \times U}$, where $\ell = 0, \ldots, W-1$ are the channel "taps" in the TD. We first introduce the model used to characterize the presence of antenna correlation at the BS side[2] which occurs in the TD. Specifically, we use the standard correlation model from [57] and express the $\ell$th TD channel matrix as follows:

$$\widehat{\mathbf{H}}_\ell = \mathbf{R}^{1/2} \widehat{\mathbf{H}}_\ell^{\text{uncor}}. \tag{6.2}$$

Here, $\widehat{\mathbf{H}}_\ell^{\text{uncor}}$ represents an uncorrelated TD channel matrix and $\mathbf{R} \in \mathbb{C}^{B \times B}$ is a correlation matrix that contains ones on the main diagonal and $\delta \in \mathbb{R}$ on the

[2] In massive MU-MIMO systems, the UEs signals are likely uncorrelated as they are spatially well-separated over potentially large cells or UE scheduling avoids correlated UEs; in contrast, the antennas at the BS are typically confined to a small area, which increases the potential for receive-side correlation.

off-diagonals. We allow $\delta$ to be either real positive or negative as long as $\delta^2 \leq 1$. We rewrite the correlation matrix as $\mathbf{R} = (1-\delta)\mathbf{I}_B + \delta \mathbf{1}_B$, where $\mathbf{I}_B$ and $\mathbf{1}_B$ is the $B \times B$ identity and all-ones matrix, respectively. We note that in the absence of receive-side correlation, i.e., $\delta = 0$, we have $\mathbf{R} = \mathbf{I}_B$ and $\widehat{\mathbf{H}}_\ell = \widehat{\mathbf{H}}_\ell^{\text{uncor}}$.

In order to take into account the practically-relevant case of imperfect CSI at the BS, we assume that the FD channel matrices $\mathbf{H}_\omega$, $\omega = 0, \ldots, W-1$, are obtained from the TD matrices $\widehat{\mathbf{H}}_\ell \in \mathbb{C}^{B \times U}$, $\ell = 0, \ldots, W-1$, via the discrete Fourier transform [169] as follows:[3]

$$\mathbf{H}_\omega = \sum_{\ell=0}^{L-1} \widehat{\mathbf{H}}_\ell \exp\left(-\frac{j2\pi\omega\ell}{W}\right) + \sigma \mathbf{E}_\omega, \tag{6.3}$$

where the matrix $\mathbf{E}_\omega \in \mathbb{C}^{B \times U}$ models channel estimation error on subcarrier $\omega$ and the parameter $\sigma \in \mathbb{R}^+$ determines the intensity of channel-estimation errors; $\sigma = 0$ corresponds to the case for perfect CSI. We assume that the entries of the matrix $\mathbf{E}_\omega$ are i.i.d. (across entries and subcarriers) circularly-symmetric complex Gaussian with unit variance. Equation (6.3) relies on the assumption that at most $L \leq W$ of the first channel taps are non-zero (or dominant) and the remaining ones are zero (or insignificant), i.e., $\widehat{\mathbf{H}}_\ell = \mathbf{0}_{B \times U}$ for $\ell = L, \ldots, W-1$. In practical OFDM and SC-FDMA systems, the maximum number of non-zero channel taps should not exceed the cyclic prefix length (assuming perfect synchronization). Hence, we can safely assume that $L \ll W$ in practical scenarios and for most standards, such as IEEE 802.11n [47] or 3GPP-LTE [159].

[3] One could improve the channel estimates $\mathbf{H}_\omega$ by exploiting the fact that only the first $L$ taps of $\mathbf{H}_\omega$ are active in the TD. An analysis of such channel estimation algorithms is left for future work.

### 6.2.3 Linear and Non-linear Data Detection and Precoding

In the massive MU-MIMO uplink, linear and non-linear data detection methods were shown to achieve near-optimal error-rate performance [5, 39]. For linear minimum mean-square error (MMSE) equalization, one first computes the Gram matrix $\mathbf{G}_\omega = \mathbf{H}_\omega^{\mathrm{H}}\mathbf{H}_\omega$ and then, computes an estimate of the transmit vector as $\hat{\mathbf{s}}_\omega = (\mathbf{G}_\omega + \mathbf{I}\frac{N_0}{E_s})^{-1}\mathbf{H}_\omega^{\mathrm{H}}\mathbf{y}_\omega$, where $N_0$ and $E_s$ stand for the noise variance and average energy per transmit symbol, respectively. For non-linear data detectors, such as the one in [39], one can operate directly on the Gram matrix $\mathbf{G}_\omega$ and the matched filter $\mathbf{H}_\omega^{\mathrm{H}}\mathbf{y}_\omega$ without any performance loss [149].

For such algorithms, a direct computation of the Gram matrix $\mathbf{G}_\omega$ for every subcarrier results in excessively high complexity. In fact, even by exploiting symmetries, $2BU^2$ real-valued multiplications are required, which is more than 16 k multiplications per subcarrier for a system with 128 BS antennas and 8 users. Furthermore, the hardware design for massive MU-MIMO data detection in [22] confirms this observation and shows that computing the Gram matrix dominates the overall hardware complexity and power consumption by at least $2\times$.

In the downlink, ZF of Wiener filter precoding are most commonly used [5]. For example, ZF precoding computes $\mathbf{x}_\omega = \mathbf{H}_\omega^{\mathrm{H}}\mathbf{G}_\omega^{-1}\mathbf{s}_\omega$, where $\mathbf{x}_\omega$ is the $B$-dimensional transmit signal and $\mathbf{s}_\omega$ the data vector. If the Gram matrix $\mathbf{G}_\omega$ has been precomputed for equalization in the uplink phase, then it can be re-used for ZF precoding in the downlink to minimize recurrent operations. Hence, to minimize the overall complexity of equalization and precoding, efficient ways to compute the Gram matrix $\mathbf{G}_\omega$ on all active subcarriers $\omega \in \Omega$ are required.

## 6.3 Interpolation-based Gram Matrix Computation

We now discuss exact and approximate interpolation-based methods for low-complexity Gram matrix computation. We note that the exact Gram matrix interpolation assumes that we have perfect CSI, so throughout this section, we will assume that $\sigma = 0$. In Section 6.4, however, we will relax the perfect-CSI assumption and study the performance of exact and approximate interpolation methods with imperfect CSI.

As a result of (6.3), the Gram matrices in the FD are given by

$$\mathbf{G}_\omega = \sum_{\ell=0}^{L-1} \sum_{\ell'=0}^{L-1} \widehat{\mathbf{H}}_\ell^{\mathrm{H}} \widehat{\mathbf{H}}_{\ell'} \exp\left( \frac{j2\pi\omega(\ell - \ell')}{W} \right) \tag{6.4}$$

for $w = 0, 1, \ldots, W-1$. Given the FD channel matrices $\mathbf{H}_\omega$ for all active subcarriers $\omega \in \Omega$, a straightforward "brute-force" approach simply computes $\mathbf{G}_\omega = \mathbf{H}_\omega^{\mathrm{H}} \mathbf{H}_\omega$ for each active subcarrier $\omega \in \Omega$. In order to reduce the complexity of such a brute-force approach, we next discuss exact and approximate Gram-matrix interpolation methods that take advantage of the facts that (i) the channel matrices (and hence, the Gram matrices) are "smooth" (correlated) across subcarriers if $L < W$ and (ii) massive MU-MIMO benefits from the well-known channel hardening effect [5, 53].

### 6.3.1 Exact Gram-Matrix Interpolation

The Gram matrix $\mathbf{G}_\omega$ in (6.4) is a Laurent polynomial matrix in the variable $x_\omega = \exp(j2\pi\omega/W)$; we refer the reader to [27] for more details on Laurent polynomial matrices. Hence, we can establish the following result for exact Gram-matrix interpolation; a short proof is given in Appendix A.4.1.

**Lemma 36.** *The Gram matrices $\mathbf{G}_\omega$ in (6.4) for all subcarriers $\omega = 0, \ldots, W$ are fully determined by $2L-1$ distinct and non-zero Gram-matrix base-points.*

Consequently, one can interpolate all of the Gram matrices *exactly* from only $2L-1$ distinct and non-zero Gram-matrix base-points that have been computed explicitly.

In order to perform exact interpolation, we first define a set of base points $\mathcal{P} \subset \Omega$ that contains $|\mathcal{P}| \geq 2L-1$ distinct subcarrier indices. We denote the $k$th base-point index as $p_k$, where $k = 0, \ldots, |\mathcal{P}|-1$, and the set of all base-point indices as $\mathcal{P} = \{p_0, \ldots, p_{|\mathcal{P}|-1}\}$. For each subcarrier index in the base-point set $\mathcal{P}$, we then explicitly compute $|\mathcal{P}| \geq 2L-1$ Gram matrices $\mathbf{G}_\omega = \mathbf{H}_\omega^{\mathrm{H}} \mathbf{H}_\omega$, $\omega \in \mathcal{P}$, and perform entry-wise interpolation for the gram matrices $\mathbf{G}_\omega$ on all remaining active subcarriers $\omega \in \Omega \backslash \mathcal{P}$.

The exact interpolation procedure for each entry is as follows. For a fixed entry $(m,n)$, we define the vector $\mathbf{g}_\mathcal{P} \in \mathbb{C}^{|\mathcal{P}|}$, which is constructed from the $(m,n)$ entries $[\mathbf{G}_\omega]_{m,n}$ taken from base-points $\omega \in \mathcal{P}$, i.e., $\mathbf{g}_\mathcal{P} = \left[ [\mathbf{G}_{p_0}]_{m,n} \cdots [\mathbf{G}_{p_{|\mathcal{P}|-1}}]_{m,n} \right]^{\mathrm{T}}$. Then, the vector $\mathbf{g}_{\Omega\backslash\mathcal{P}} \in \mathbb{C}^{|\Omega|-|\mathcal{P}|}$ that contains the entry $(m,n)$ for all remaining Gram matrices $\mathbf{G}_\omega$, $\omega \in \Omega \backslash \mathcal{P}$, is given by

$$\mathbf{g}_{\Omega\backslash\mathcal{P}} = \mathbf{F}_{\Omega\backslash\mathcal{P},L} (\mathbf{F}_{\mathcal{P},L}^{\dagger} \mathbf{g}_\mathcal{P}). \tag{6.5}$$

Here, $\mathbf{F}_{\mathcal{P},L}$ represents a $|\mathcal{P}| \times (2L-1)$ matrix where we take the $|\mathcal{P}|$ rows indexed by $\mathcal{P}$ and the first $L$ and last $L-1$ columns from the $W$-point discrete Fourier transform (DFT) matrix; the entries of the DFT matrix are defined as $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{W}} \exp\Big(-\frac{j2\pi}{W}(m-1)(n-1)\Big)$. Similarly, $\mathbf{F}_{\Omega\backslash\mathcal{P},L}$ represents a $(|\Omega|-|\mathcal{P}|) \times (2L-1)$ matrix where we take $|\Omega|-|\mathcal{P}|$ rows indexed by $\Omega\backslash\mathcal{P}$ and the first $L$ and last $L-1$ columns from the $W$-point DFT matrix. In words,

the exact interpolation method in (6.5) first computes the $2L-1$ TD Gram-matrix entries and then, transforms these elements into the frequency domain via the DFT. See [24–27] for additional details on other exact interpolation methods developed for MIMO systems.

Although the method in (6.5) is able to *exactly* interpolate the Gram matrix across all $W$ subcarriers, it is in many situations not practical due to the high complexity of the matrix inversion required in $\mathbf{F}_{\mathcal{P},L}^{\dagger}\mathbf{g}_{\mathcal{P}}$. If, however, one can sample the base points uniformly over all $W$ tones, the complexity of matrix inversion can be reduced significantly. Unfortunately, this approach is often infeasible in practice due to the presence of guard-band constraints in OFDM-based or SC-FDMA-based standards [47, 159]. Another issue of exact interpolation methods, such as the ones in [24–27] and ours in (6.5), is that they generally assume perfect CSI and no BS-antenna correlation. As we will show in Section 6.6.3, imperfect CSI results in poor interpolation performance—this is due to the fact that the matrix $\mathbf{F}_{\mathcal{P},L}$ is typically ill-conditioned, especially when sampling Gram-matrices close to the minimum number of $2L-1$ base points.

We next propose two approximate interpolation schemes that not only require (often significantly) lower complexity than a brute-force approach or exact interpolation in (6.5), but also approach the performance of a brute-force approach in massive MU-MIMO systems and are robust to channel-estimation errors.

### 6.3.2 Approximate Gram-Matrix Interpolation

We consider the following two approximate Gram-matrix interpolation methods illustrated in Fig. 6.1.

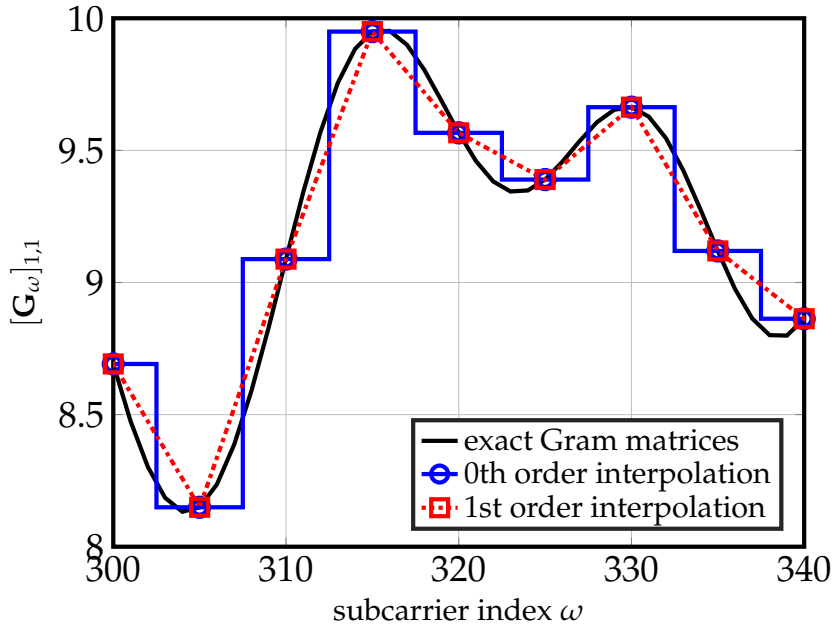



Figure 6.1: Illustration of 0th and 1st order interpolation for the entry $[\mathbf{G}_\omega]_{1,1}$ across subcarriers. We explicitly compute the Gram matrix for every fifth subcarrier index $\omega$ and interpolate the remaining matrices. We assume a 128 BS antenna, 8 user massive MU-MIMO system with $W = 2048$ subcarriers, a delay spread of $L = 144$, and perfect CSI without BS-antenna correlation.

**0th Order Interpolation**

We select a set of $|\mathcal{P}|$ distinct base-points with $\mathcal{P} = \{p_0, \ldots, p_{|\mathcal{P}|-1}\} \subset \Omega$. We explicitly compute $\mathbf{G}_p = \mathbf{H}_p^{\mathrm{H}}\mathbf{H}_p$ on these base points and perform 0th order (or nearest-neighbor) interpolation for the remaining subcarriers in the set $\Omega \backslash \mathcal{P}$ according to:

$$\widetilde{\mathbf{G}}_\omega = \mathbf{G}_p, \quad p = \arg\min_{\tilde{p} \in \mathcal{P}} |\tilde{p} - \omega|, \quad \forall \omega \in \Omega \backslash \mathcal{P}. \tag{6.6}$$

In words, we set the interpolated Gram matrix $\widetilde{\mathbf{G}}_\omega$ equal to the nearest Gram matrix that has been computed explicitly for one of the neighboring base points.

**1st Order Interpolation**

Analogously to the 0th order interpolation method, we explicitly compute $\mathbf{G}_p = \mathbf{H}_p^{\mathrm{H}}\mathbf{H}_p$ on a selected set of base-points $p \in \mathcal{P}$. Then, for each target subcarrier $\omega \in \Omega\backslash\mathcal{P}$ we pick two nearest base-points $p_k$ and $p_{k+1}$, i.e., $p_k \leq \omega \leq p_{k+1}$, and perform entry-wise linear interpolation according to

$$\widetilde{\mathbf{G}}_\omega = \lambda_\omega \mathbf{G}_{p_k} + (1-\lambda_\omega)\mathbf{G}_{p_{k+1}}, \quad \omega \in \Omega\backslash\mathcal{P}, \tag{6.7}$$

where $\lambda_\omega = (p_{k+1}-\omega)/(p_{k+1}-p_k)$ and $p_k < p_{k+1}$.

## 6.4 Approximation Error Analysis

We now analyze the approximation error associated with the approximate interpolation schemes from Section 6.3.2. We use $\mathbf{G}_\omega$ to represent the Gram matrices that have been computed exactly and $\widetilde{\mathbf{G}}_\omega$ to represent the Gram matrices that are obtained via approximate interpolation. Evidently, the exact interpolation scheme in Section 6.3.1 entails no approximation error.

### 6.4.1 Mean-Square-Error of Approximate Interpolation

We study the mean-squared error (MSE) on each entry $(m,n)$ for the $\omega$-th subcarrier, which we define as follows:

$$o\text{-}\mathsf{MSE}_\omega^{(m,n)} \triangleq \mathbb{E}\Big[\big|[\widetilde{\mathbf{G}}_\omega]_{m,n} - [\mathbf{G}_\omega]_{m,n}\big|^2\Big]. \tag{6.8}$$

Here, $o$ represents the order of interpolation, i.e., we have either $o = 0$ or $o = 1$. Our results make extensive use of the scaled Fejér kernel [170] given by

$$f_L(\phi) = L^{-2}\frac{1-\cos(L\phi)}{1-\cos(\phi)} = L^{-2}\frac{\sin^2(L\phi/2)}{\sin^2(\phi/2)} \tag{6.9}$$

and rely on the following key properties of this kernel; the proof is given in Appendix A.4.2.

**Lemma 37.** *The scaled Fejér kernel (6.9) is non-negative, bounded from above by one, and monotonically decreasing in $\phi$ for $\phi \in [0, 2\pi/L]$ with $L > 1$.*

### 6.4.2 Mean-Square-Error of 0th Order Interpolation

The following result precisely characterizes the MSE of 0th order interpolation for imperfect CSI as in (6.3) and BS-antenna correlation as in (6.2); the proof is given in Appendix A.4.3.

**Theorem 38.** *Let the entries of the TD matrices $\widehat{\mathbf{H}}_\ell^{\text{uncor}}$, $\ell = 1, \ldots, L$, be distributed $\mathcal{CN}(0, 1/(BL))$ per complex entry. Assume that the off-diagonal of the receive correlation matrix be $\delta$, the variance of the channel estimation error to be $\sigma$, and $p \in \mathcal{P}$ is the closest base point to the target subcarrier $\omega$. Then, for any $(m, n)$ entry of the Gram matrix $\mathbf{G}_\omega$, the MSE for the 0th order interpolation method in (6.6) is given by*

$$\text{0-MSE}_\omega = \varepsilon_{CSI} + \frac{2}{B}(1+\varepsilon_{cor})\left(1 - f_L\left(\frac{2\pi}{W}(p-\omega)\right)\right), \tag{6.10}$$

*where we use the definitions*

$$\varepsilon_{CSI} = 2\sigma^2(2 + B\sigma^2) \quad \textit{and} \quad \varepsilon_{cor} = \delta^2(B-1).$$

From this result, we observe that, as the number of BS antennas $B$ increases, $\varepsilon_{\text{CSI}}$ increases quadratically with respect to $\sigma^2$. For perfect CSI, i.e., $\sigma = 0$ so

$\varepsilon_{\text{CSI}} = 0$, the MSE for 0th order interpolation decreases with an increasing number of BS antennas $B$ as $\frac{\partial}{\partial B}$0-MSE$_\omega < 0$, if $\delta^2 < 1$. Also, we note in the case for non-zero correlation, i.e., $\delta \neq 0$, the MSE for 0th order interpolation is amplified (compared to that with no correlation) by a factor of $1 + \delta^2(B-1)$. Furthermore, we observe that the MSE is independent of the entry of the Gram matrix (i.e., the MSE is identical for the diagonal as well as off-diagonal entries); this is a consequence of the i.i.d. assumption of the TD channel matrices $\widehat{\mathbf{H}}_\ell^{\text{uncor}}$.

To gain additional insight into the behavior of 0th order interpolation in the large BS-antenna limit, we have the following result.

**Corollary 39.** *Assume the conditions in Theorem 38, and let $\sigma^2 \to 0$. Then, as $B \to \infty$, the MSE of 0th order interpolation is given by*

$$\lim_{B\to\infty} \text{0-MSE}_\omega = 2\delta^2\left(1 - f_L\left(\frac{2\pi}{W}(p-\omega)\right)\right).$$

Corollary 39 demonstrates that in the large-BS antenna limit, the MSE of 0th order interpolation is zero across all subcarriers if and only if the BS antennas are uncorrelated, i.e., $\delta = 0$. For $\delta \neq 0$, the MSE depends on the distance between the nearest base-point and the target subcarrier.

While the MSE expression in (6.10) is exact, it does not provide much intuition. We define the following quantity that enables us to further analyze the MSE in (6.10).

**Definition 16.** *The maximum distance between any subcarrier $\omega$ and the nearest base point is given by:*

$$d_{\max} = \max_{k\in\mathcal{P}} \left\lfloor \frac{p_{k+1} - p_k}{2} \right\rfloor. \tag{6.11}$$

With the maximum distance $d_{\max}$ for a given set of base points $\mathcal{P}$, Corollary 40 shows that the 0th order approximation can be bounded from above using simple analytic expressions; the proof is given in Appendix A.4.4.

**Corollary 40.** *Let $d_{\max}$ be the maximum distance in (6.11) and assume the conditions in Theorem 38 hold. Then, the maximum MSE of 0th order interpolation over all active subcarriers $\omega \in \Omega$ is bounded by:*

$$\max_{\omega\in\Omega}\{\text{0-MSE}_\omega\} \leq \begin{cases} \varepsilon_{CSI} + \dfrac{2}{B}(1+\varepsilon_{cor}), & d_{\max} \geq \dfrac{W}{L} \\ \varepsilon_{CSI} + \dfrac{2(1+\varepsilon_{cor})}{B}\left(1 - f_L\left(\dfrac{2\pi}{W} d_{\max}\right)\right), & d_{\max} < \dfrac{W}{L}. \end{cases}$$

Corollary 40 implies that regardless of small or large maximum distance $d_{\max}$, the MSE given by the 0th order approximation always decreases with the number of BS antennas $B$ if $\delta = 0$ and $\sigma^2 = 0$ (see also Theorem 38). In addition, if the distance between the interpolated subcarrier index $\omega$ and its closest base point is sufficiently small, i.e., $d_{\max} < W/L$, then we obtain a sharper upper bound on the MSE than $\varepsilon_{\text{CSI}} + \frac{2}{B}(1+\varepsilon_{\text{cor}})$. In a scenario with a large delay spread $L$, Corollary 40 reveals that one requires finer-spaced base points for 0th order interpolation in order to keep the approximation error strictly smaller than $\varepsilon_{\text{CSI}} + \frac{2}{B}(1+\varepsilon_{\text{cor}})$. Since the maximum error is mainly determined by $d_{\max}$, a good strategy for selecting base points with 0th order approximation is uniformly spacing them in the set of active subcarriers $\Omega$.

### 6.4.3 Mean-Square-Error of 1st Order Interpolation

We now present the approximation error analysis of 1st order interpolation. The following result characterizes the MSE of 1st order interpolation; the proof is given in Appendix A.4.5.

**Theorem 41.** *Let the entries of the TD matrices $\widehat{\mathbf{H}}_{\ell}^{\text{uncor}}$, $\ell = 1, \ldots, L$, be distributed $\mathcal{CN}(0, 1/(BL))$ per complex entry. Assume that the off-diagonal of the receive correlation matrix be $\delta$, the variance of the channel estimation error to be $\sigma$ across all subcarriers $\omega = 0, \ldots,$, and $p \in \mathcal{P}$ is the closest base point to the target subcarrier $\omega$. Then, for any $(m, n)$-th entry of the Gram matrix $\mathbf{G}_{\omega}$, the MSE for the 1st order interpolation method in (6.6) is given by*

$$\begin{aligned} \text{1-MSE}_{\omega} = \varepsilon_{CSI}(1 - \lambda_{\omega}(1 - \lambda_{\omega})) \\ + \frac{2}{B}(1 + \varepsilon_{cor})\big(1 - \lambda_{\omega}(1 - \lambda_{\omega}) + \lambda_{\omega}(1 - \lambda_{\omega}) f_L(\theta) \\ - (1 - \lambda_{\omega}) f_L(\lambda_{\omega}\theta) - \lambda_{\omega} f_L((1 - \lambda_{\omega})\theta)\big), \end{aligned} \tag{6.12}$$

*where $\theta = \frac{2\pi}{W}(p_{k+1} - p_k)$ and $\lambda_{\omega} = (p_{k+1} - \omega)/(p_{k+1} - p_k)$.*

Analogously to 0th order interpolation, we observe that the MSE of 1st order interpolation is independent of the entry $(m, n)$ and impacted by CSI errors and receive correlation (see Section 6.4.2 for detailed discussion). The result shown next in Corollary 42 reveals that if the spacing between the two base-points $p_k$ and $p_{k+1}$ defined as $d_k = p_{k+1} - p_k$ is sufficiently small, then the 1st order interpolation strictly outperforms 0th order interpolation, i.e., $\text{1-MSE}_{\omega} < \text{0-MSE}_{\omega}$ for all $\omega \in (p_k, p_{k+1})$; the proof is given in Appendix A.4.6.

**Corollary 42.** *Let $d_k$ denote the spacing between two base-points $p_k$ and $p_{k+1}$, and assume the conditions in Theorem 38 and Theorem 41 hold. If $d_k \leq W/(3L)$, then*

$$\text{1-MSE}_{\omega} \leq \text{0-MSE}_{\omega}, \quad \text{for all } \omega \in (p_k, p_{k+1}), \tag{6.13}$$

*which holds with equality if and only if $L = 1$ and $\varepsilon_{CSI} = 0$.*

We note that the condition $d_k < W/(3L)$ is not sharp; Appendix A.4.6 outlines the details on how it can be sharpened. Furthermore, given that $d_k$ is significantly

larger than $W/(3L)$, we can construct situations for which 0th order interpolation *outperforms* 1st order interpolation. Note that for $L = 1$, the FD channel is flat (i.e., $\mathbf{G}_\omega$ is constant for all $\omega$) and hence, 1st and 0th order interpolation have the same MSE.

In summary, we observe that for both approximate interpolation methods, the MSE can be lowered by increasing the number of BS antennas $B$ assuming that the channel estimation error $\varepsilon_{\text{CSI}}$ decreases with $B$. In the large-antenna limit $B \to \infty$ with perfect CSI and no BS-antenna correlation, the MSE vanishes, which is an immediate consequence of channel hardening in massive MU-MIMO systems [5]. Furthermore, 1st order interpolation generally outperforms 0th order interpolation for a sufficiently small minimum spacing between adjacent base points, i.e., for $d_k \leq W/(3L)$.

## 6.5 Complexity Analysis

We next compare the computational complexity of the four studied Gram-matrix computation algorithms: brute-force computation, exact interpolation, 0th order interpolation, and 1st order interpolation. We measure the computational complexity by counting the number of real-valued multiplications.[4]

[4]We assume that a complex-valued multiplication requires four real-valued multiplications; computation of the squared magnitude of a complex number is assumed to require two real-valued multiplications.

### 6.5.1 Brute-Force Computation

We start by deriving the total computational complexity required by the brute-force (BF) method. We only compute the upper triangular part of $\mathbf{G}_\omega$ (since the matrix is Hermitian). Each off-diagonal entry requires $B$ complex-valued multiplications, which corresponds to $4B$ real-valued multiplications; each diagonal entry requires only $2B$ real-valued multiplications. Hence, the computational complexity of computing $\mathbf{G}_\omega$ using the BF method is

$$C_{\text{BF}} = |\Omega| \left( 4B \frac{U(U-1)}{2} + 2BU \right) = 2|\Omega| BU^2 \tag{6.14}$$

for a total number of $|\Omega|$ active subcarriers.

### 6.5.2 Exact Interpolation

We now derive the computational complexity of exact interpolation as discussed in Section 6.3.1. Exact interpolation requires a BF computation of the Gram matrix at each of the $|\mathcal{P}|$ base points. We will use the $|\mathcal{P}|$ precomputed base points of $\mathbf{G}_\omega$ to interpolate the remaining $|\Omega| - |\mathcal{P}|$ Gram matrices.

We will assume that the base points and the assumed channel delay spread $L$ are fixed *a-priori* so that $\mathbf{F}_{\Omega\setminus\mathcal{P},L}\mathbf{F}_{\mathcal{P},L}^\dagger$ in (6.5) can be precomputed and stored. We emphasize that this approach does not include the computational complexity of computing the interpolation matrix itself, which favors this particular interpolation scheme from a complexity perspective. In fact, we only need to multiply the precomputed interpolation matrix $\mathbf{F}_{\Omega\setminus\mathcal{P},L}\mathbf{F}_{\mathcal{P},L}^\dagger$ with the vector $\mathbf{g}_\mathcal{P}$, which requires $4(|\Omega| - |\mathcal{P}|)|\mathcal{P}|$ real-valued multiplications. Hence, the total

computational complexity of exact interpolation is:

$$\begin{aligned} C_{\text{Exact}} &= \frac{|\mathcal{P}|}{|\Omega|} C_{\text{BF}} + 4(|\Omega| - |\mathcal{P}|)|\mathcal{P}| \frac{U(U+1)}{2} \\ &= 2|\mathcal{P}|(|\Omega| - |\mathcal{P}| + B)U^2 + 2|\mathcal{P}|(|\Omega| - |\mathcal{P}|)U. \end{aligned} \tag{6.15}$$

We note that if the number of users $U$ is large and the number of base points is similar to the number of BS antennas, i.e., $|\mathcal{P}| \simeq B$, then the BF method in (6.14) and exact interpolation (6.15) exhibit similar complexity. We also observe that the complexity of exact interpolation (6.15) is lower than that of the BF method (6.14) if $|\mathcal{P}| < (1+U)^{-1}BU$. Since the use of $|\mathcal{P}| \geq 2L-1$ distinct base points guarantees exact interpolation (assuming perfect CSI), we observe that exact interpolation has lower complexity than the BF method if $L$ is (approximately) smaller than $B/2$.

### 6.5.3 0th Order Interpolation

The computational complexity of the 0th order interpolation method is given by

$$C_{\text{0th}} = 2|\mathcal{P}|BU^2, \tag{6.16}$$

as we only need to compute the Gram matrices on all the base points. We note that since typically $|\mathcal{P}| \ll |\Omega|$ the savings (in terms of real-valued multiplications) are significant compared to the BF approach and exact interpolation, but does so at the cost of approximation errors (cf. Section 6.6.4).

### 6.5.4 1st Order Interpolation

The computational complexity of the 1st order interpolation is given by

$$\begin{aligned} C_{\text{1st}} &= C_{\text{0th}} + 4(|\Omega| - |\mathcal{P}|)\frac{U(U+1)}{2} \\ &= 2|\mathcal{P}|BU^2 + 2(|\Omega| - |\mathcal{P}|)U(U+1), \end{aligned} \tag{6.17}$$

where we assume that the interpolation weight $\lambda_\omega$ was precomputed. We note that the linear interpolation stage for each subcarrier $\omega \in \Omega \backslash \mathcal{P}$ requires four real-valued multiplications. By comparing (6.16) to (6.17), we observe that the complexity of 1st order interpolation always exceeds the complexity of the 0th order method, but the complexity is significantly lower than that of the BF method as we generally have $|\mathcal{P}| \ll |\Omega|$.

## 6.6 Numerical Results

We now study the MSE, the error-rate performance, and the computational complexity of the proposed Gram-matrix interpolation schemes. We consider a MU-MIMO-OFDM system with 128 BS antennas and with 8 single-antenna users. We assume a total of $W = 2048$ subcarriers, with $|\Omega| = 1200$ active subcarriers, similar to that used in 3GPP LTE [159]. Unless stated otherwise, we assume that the entries of the TD channel matrices are i.i.d. circularly-symmetric complex Gaussian with variance $1/(BL)$ and we consider 16-QAM transmission (with Gray mapping). We use a linear MMSE equalizer for data detection; see Section 6.2.3. For situations with imperfect CSI, we consider pilot-based maximum-likelihood (ML) channel estimation with a single orthogonal pilot sequence of length $U$ with the same transmit power as for the data symbols.

### 6.6.1 Complexity Comparison

We now assess the complexity of the various Gram-matrix computation methods in comparison to the overall complexity required for linear MMSE-based data detection, which includes Gram-matrix and matched-filter computation as well as matrix inversion for each active subcarrier. The results shown here are for a $128 \times 8$ (the notation represents $B \times U$) massive MU-OFDM-MIMO system with 1200 active subcarriers and a delay spread of $L = 144$.

Fig. 6.2 compares the complexity of Gram matrix computation for four different methods, brute-force, exact, 0th-, and 1st-order interpolation methods for $|\mathcal{P}| = \{0.25|\Omega|, 0.5|\Omega|, 0.75|\Omega|, |\Omega|\}$. The solid part of the bar plot shows the complexity of Gram matrix computation; the fenced part corresponds to the remaining complexity required for data detection (including matched-filter computation and a matrix inversion for each active subcarrier). The percentage values indicate the relative complexity of Gram-matrix computation compared to the total complexity required for data detection. We assume a Cholesky-based implicit matrix inversion for detection [171]. As demonstrated in [139], Gram matrix computation requires majority of the computational complexity, as it scales quadratically in the number of BS antennas.

We see that the exact interpolation method results in high complexity in the considered system (see Section 6.5.2 for exact details when exact interpolation achieves lower complexity than a BF approach). We also see that the proposed 0th and 1st order approximation methods both achieve significant complexity reductions. For $|\mathcal{P}| = 0.25|\Omega|$, the proposed methods requires less than half the complexity of a BF approach. As we will show in Section 6.6.3, the proposed approximate interpolation methods will exhibit similar error-rate

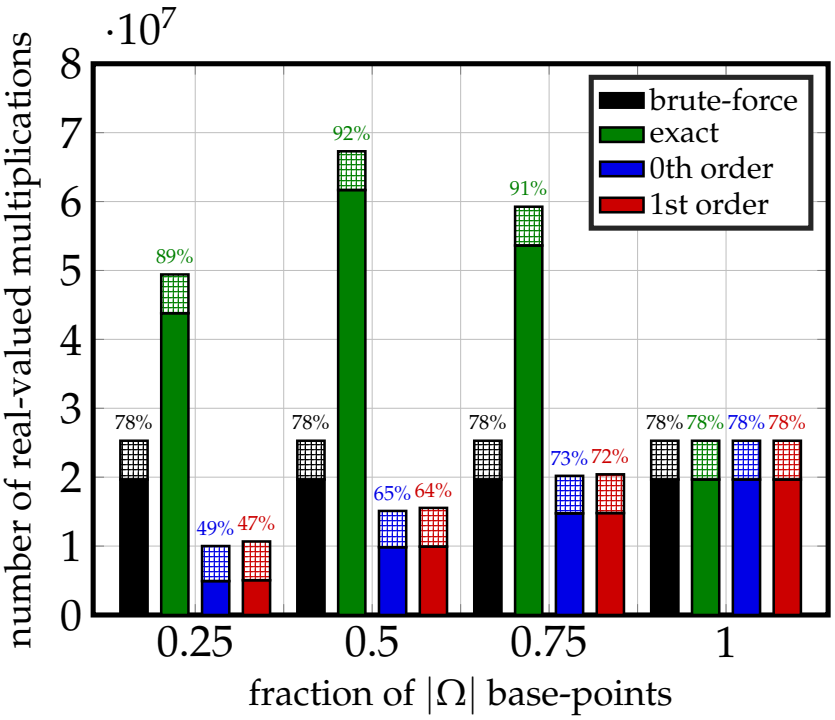



Figure 6.2: Comparison of Gram matrix computational complexity for exact and approximate interpolation, and brute-force methods. We compare the complexity of various Gram matrix computation methods in comparison to the complexity of data detection for a $128 \times 8$ MU-OFDM-MIMO system with $|\Omega| = 1200$, $L = 144$, and four different sets of base-points $|\mathcal{P}| = \{0.25|\Omega|, 0.5|\Omega|, 0.75|\Omega|, |\Omega|\}$. The percentage values shown in the above bar plots show the relative percentage of Gram-matrix computation compared to the total complexity required for data detection.

performance as that the BF approach (see Figs. 6.4–6.7), but does so at fraction of the computational complexity.

### 6.6.2 MSE of Approximate Interpolation

Fig. 6.3 compares the MSE of 0th and 1st order interpolation as proposed in Section 6.3.2. Note that the BF method and exact interpolation have an MSE of zero and hence, we exclude these results. We select two base points at subcarriers 500 and 600 and one target point at subcarrier 512, and compare the MSE for

different numbers of BS antennas and under ideal and non-ideal scenarios. In the ideal scenario, we assume perfect CSI and no BS-antenna correlation, whereas in the non-ideal scenario we assume channel-estimation at $\mathsf{SNR} = 25$ dB across all subcarriers with the signal-to-noise ratio (SNR) defined by $\mathsf{SNR} = U/(B\sigma^2)$ and a BS-antenna correlation of $\delta = 0.1$. In order to assess the approximation error with respect to different channel delay spreads, we set $L \in \{36, 72, 144\}$. The resulting MSE is shown in Fig. 6.3. Note that the MSE for both 0th and 1st are independent of the entry (as predicted by Theorems 38 and 41); hence, we consider the average MSE across all entries.

We observe that the 1st order interpolation method achieves a lower MSE than that given by 0th order interpolation, where the performance gap increases with larger delay spreads $L$. This observation is caused by the fact that for small delay spreads $L$, the channel is more smooth across subcarriers. For larger delay spreads $L$, 1st order interpolation captures the faster-changing behavior of the Gram matrix, whereas the 0th order interpolation ignores such changes. We also see that the MSE degrades in the non-ideal scenario, even if we increase the number of BS antennas; this behavior is reflected in our analytical results. Finally, we see that the simulated MSE matches perfectly our theoretical results in Theorems 38 and 41.

### 6.6.3 Error-rate Performance

We now compare the error-rate performance of the proposed Gram-matrix computation schemes. We simulate the bit-error rate (BER) for a MU-MIMO-OFDM system for a different number of base-points $|\mathcal{P}|$ and for perfect as well as im-

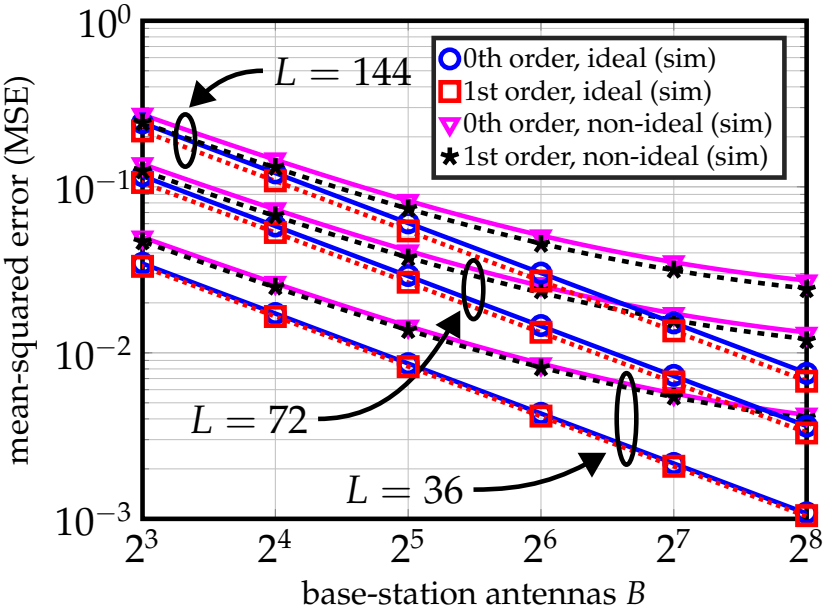



Figure 6.3: MSE of 0th and 1st order interpolation for an entry of $\mathbf{G}_{512}$ using two base points $\mathbf{G}_{500}$ and $\mathbf{G}_{600}$ for three different delay spreads $L \in \{36, 72, 144\}$ for the ideal and non-ideal scenarios. The markers represent simulation results whereas the lines represent our approximation-error analysis. Evidently, our theory matches perfectly with the simulated values.

perfect CSI. We also investigate the impact of a more realistic channel model. For all results, we simulate three different numbers of base-points $|\mathcal{P}| = L/4$, $|\mathcal{P}| = 2L - 1$, and $|\mathcal{P}| = 4L$, and select equally-spaced base points. Figures 6.4 and 6.5 show BER simulation results for an i.i.d. Rayleigh fading scenario with perfect and imperfect CSI, respectively. Figure 6.6 shows BER simulation results for the QuaDRiGa channel model[5] with imperfect CSI [172]. We note that QuaDRiGa channel model includes a path-loss model for each user.

Figure 6.4 shows that exact interpolation for $|\mathcal{P}| \geq 2L - 1$ base points provides identical results as the BF method (up to machine precision) for a system with perfect CSI. For $|\mathcal{P}| = L/4$ base points, the proposed 0th and 1st order

[5]We simulate a square antenna array with a non-line-of-sight scenario with a 2 GHz carrier frequency, 20 MHz bandwidth, and 200 m distance between BS antenna and the users. Our algorithms assume $L = 144$ but the true delay spread is slightly smaller.

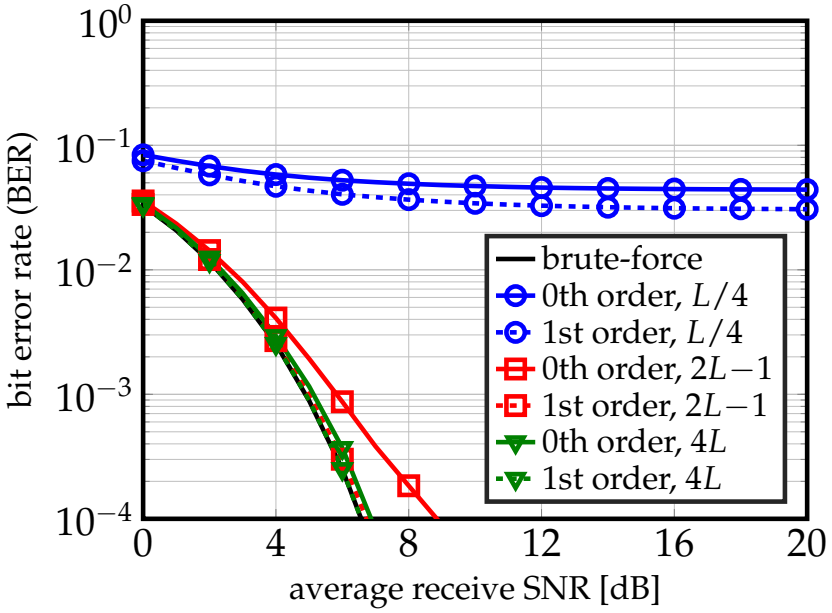



Figure 6.4: Uncoded bit error rate (BER) comparison in a 128 BS antenna, 8 (single-antenna) user, wideband massive MU-MIMO-OFDM system for an i.i.d. Rayleigh fading with perfect CSI. The values next to the legend entries correspond to the number of base points $|\mathcal{P}|$. The proposed approximation methods schemes achieves similar performance as exact brute-force method at a fraction of the complexity.

interpolation exhibit an error floor; this performance loss can be mitigated substantially by increasing the number of base points to $|\mathcal{P}| = 2L - 1$. By setting $|\mathcal{P}| = 4L < 0.5|\Omega|$, the both the 0th and 1st order interpolation methods exhibit virtually no BER performance loss.

Figure 6.5 shows the situation for imperfect CSI (with channel estimation). We observe that the performance of the BF method and that of exact interpolation are no longer equal. In fact, for $|\mathcal{P}| = 2L - 1$ and $|\mathcal{P}| = 4L$ base points, exact interpolation exhibits a significant error floor. The reason is due to the fact that the interpolation matrix is ill-conditioned, which results in significant noise enhancement artifacts. Although the error floor is decreased for $|\mathcal{P}| = 4L$, a floor remains at $10^{-3}$ BER. In contrast, the error floor of 0th order and 1st order

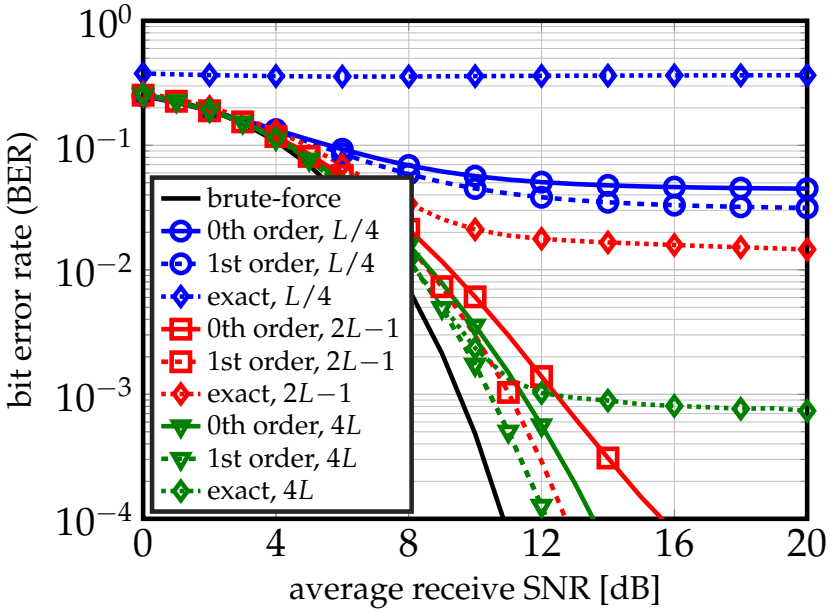



Figure 6.5: Uncoded bit error rate (BER) comparison in a 128 BS antenna, 8 (single-antenna) user, wideband massive MU-MIMO-OFDM system for an i.i.d. Rayleigh fading with channel estimation. The values next to the legend entries correspond to the number of base points $|\mathcal{P}|$. In addition, the proposed approximate interpolation schemes (0th and 1st order interpolation) outperform exact interpolation, and approach the performance of the exact brute-force method for a small number of base points.

interpolation for $|\mathcal{P}| = 2L - 1$ and $|\mathcal{P}| = 4L$ base points is well-below $10^{-4}$ BER and hence, the proposed approximate interpolation schemes are more resilient to scenarios with imperfect CSI than exact interpolation.

Figure 6.6 shows the BER performance for the QuaDRiGa channel model with imperfect CSI. We observe that all considered interpolation methods achieve a lower error floor than that given in Fig. 6.5 for $|\mathcal{P}| = L/4$; this is due to the fact that the effective delay spread for the considered channel is smaller than $L = 144$ (which is assumed in the algorithms). Once again, we observe a BER floor of exact interpolation for all considered numbers of base points. In summary, the proposed approximate interpolation methods are more robust in practical scenarios than the exact interpolation method.

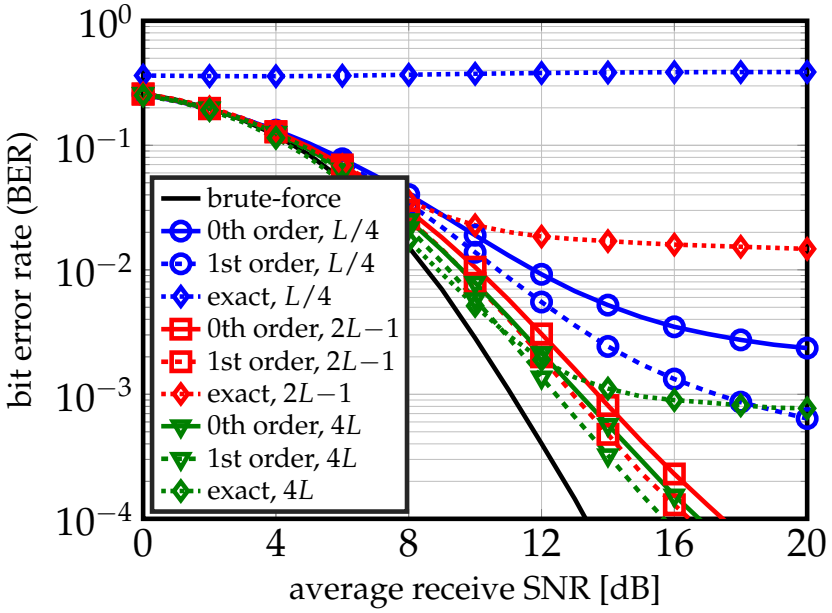



Figure 6.6: Uncoded bit error rate (BER) comparison in a 128 BS antenna, 8 (single-antenna) user, wideband massive MU-MIMO-OFDM system for a QuaDRiGa channel [172] with channel estimation. The values next to the legend entries correspond to the number of base points $|\mathcal{P}|$. In addition, the proposed approximate interpolation schemes (0th and 1st order interpolation) outperform exact interpolation, and approach the performance of the exact brute-force method for a small number of base points.

### 6.6.4 Performance and Complexity Trade-offs

We now investigate the BER performance vs. computational complexity trade-off for the proposed approximate interpolation methods with imperfect CSI. We use the complexity $C_{\text{BF}}$ of the BF method in (6.14) as our baseline, and we compare it to that of the proposed 0th and 1st order interpolation methods in (6.16) and (6.17), respectively. We vary the number of base points $|\mathcal{P}|$ from $L$ to $D$ and simulate the minimum SNR required for the linear MMSE equalizer to achieve $10^{-3}$ BER.

Figure 6.7 shows the trade-off results for 0th and 1st order interpolation. For a fixed fraction of the complexity of $C_{\text{BF}}$, we observe that the 1st order interpolation

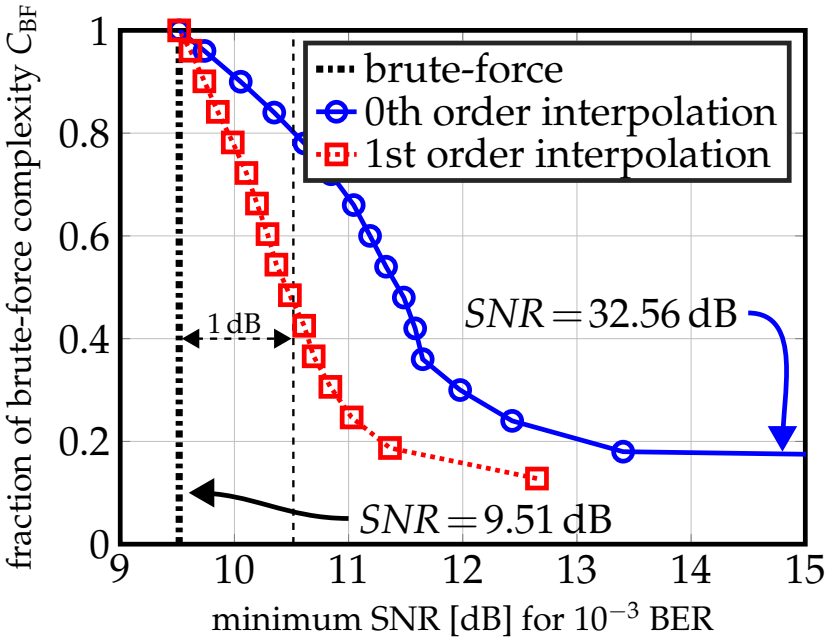



Figure 6.7: Trade-off between SNR performance and computational complexity for 0th and 1st order interpolation in an i.i.d. Rayleigh fading channel with imperfect CSI. Both approximate interpolation methods approach the performance of an exact brute-force approach at a fraction of the complexity.

method always outperforms the 0th order interpolation method. Hence, Fig. 6.7 clearly reveals that the additional complexity required by linear interpolation is beneficial when jointly considering performance and complexity. In addition, we see that the 1st order interpolation method approaches the SNR performance of the BF method by 1 dB with only 45% of the complexity.

## 6.7 Conclusions

We have studied the performance of exact and approximate interpolation-based Gram matrix computation for wideband massive MU-MIMO-OFDM systems. Instead of performing a brute-force (BF) computation of the Gram matrix for all subcarriers or using exact interpolation schemes, we have proposed two simple,

yet efficient approximate interpolation methods. We have demonstrated that channel hardening in massive MU-MIMO enables the proposed 0th and 1st order interpolation schemes to perform close to that of an exact BF computation at only a fraction of the computational complexity. In addition, the proposed approximate interpolation methods are more robust to channel-estimation errors and receive-side antenna correlation than exact interpolation methods.

CHAPTER 7

# ASIC DESIGN OF OPTIMAL DATA DETECTION FOR MASSIVE MU-MIMO

## 7.1 Introduction

Massive multi-user (MU) multiple-input multiple-output (MIMO) is believed to be a core technology for the next generation (5G) wireless systems. Massive MU-MIMO promises higher spectral efficiency, improved link reliability, and coverage over existing small-scale MIMO systems [5, 6, 140]. One critical challenge is the computational complexity of optimal data detection, which increases exponentially in the number of transmit antennas [58]. Moreover, data detection in the uplink is among the most critical processing tasks in terms of implementation complexity, power consumption, and throughput in massive MU-MIMO systems [140]. As a consequence, a successful deployment of massive MU-MIMO necessitates novel signal processing algorithms that not only provides near-optimal spectral efficiency, but also can be implemented efficiently in very-large scale integration (VLSI) circuits.

To enable high-throughput data detection performance for massive MU-MIMO systems, a variety of low-complexity algorithms (see [7, 28–30] and references therein) as well as the corresponding hardware designs in field-programmable gate array (FPGA) [8,22,31–34] and application-specific integrated circuit (ASIC) [35–37] designs have been proposed in the existing literature. Majority of these algorithms, and their hardware designs rely on (approximate) linear data detection, which are known to scale poorly in massive MU-MIMO uplink when the number of transmit antennas is similar to that of receive an-

tennas at the BS [22]. Moreover, these algorithms are designed to operate on channels that have i.i.d. Rayleigh fading, which in general, does not hold in practical massive MU-MIMO systems.

### 7.1.1 Relevant Prior Art

#### Linear and Non-linear Data Detection

While there has been plethora of data detection algorithms for massive MU-MIMO systems (see [7, 28, 29] and the references therein), very few has been integrated on FPGAs or ASICS; majority of the designs implement sub-optimal (approximate) linear methods [8, 22, 31–34]. These (approximate) linear methods have shown to provide a good compromise for performance-complexity trade-off in massive MU-MIMO systems when the UE-to-BS ratio is small, i.e., $B \gg U$. Unfortunately, (approximate) linear methods are known to perform poorly when UE-to-BS ratio increases, and also are not scalable when there are a large number of UEs, due to the cubic nature of the computation of the inverse of the regularized Gram matrix. For square systems (UE-to-BS ratio is one), reference [173] has shown excellent error-rate performance via semidefinite relaxation, however, it does not provide SISO capabilities, and is only restricted to binary (BPSK) and quadrature phase shift keying (QPSK) constellations. Unlike the aforementioned linear and non-linear data detection methods presented, our proposed LAMA algorithm is avoids the matrix inverse computation and provides significant gains in spectral efficiency compared to linear data detectors in high UE-to-BS ratios.

**Message-passing-based Data Detection**

Message-passing-based data detection algorithms have shown to be an excellent alternative to linear detection for massive MU-MIMO as they provide a superior error-rate performance at low complexity. References [35,36] have implemented [111], which demonstrated a significant spectral efficiency gain over linear detectors without having to compute an explicit matrix inverse; [111] exploits channel hardening behavior which arises in massive MU-MIMO [5]. Unfortunately, however, as shown in [38], [111] causes significant error floors for finite-dimensional systems when BS-to-UE ratio is high for higher-order constellations, and when the channel is non-i.i.d. Rayleigh and exhibits correlations, which occurs in practical massive MU-MIMO systems. To overcome the challenge in non-i.i.d. channels, [37] have implemented expectation-propagation [174]. However, expectation-propagation requires an explicit computation of the matrix inversion, which scales cubicly in the number of UEs, leading to high computational complexity.

The LAMA algorithm we propose here, and consequently the corresponding ASIC implementation, does not require an explicit computation of the matrix inverse. Unlike message-passing algorithms proposed in [111], LAMA has theoretical performance guarantees in the asymptotic limit. In addition, the LAMA algorithm borrows ideas from the generalized approximate message passing [59] and damping-techniques [124, 157] to provide a low-complexity algorithm while achieving near-optimal data detection performance for non-i.i.d. and finite-dimensional systems.

## 7.2 Massive MIMO Detection

### 7.2.1 System Model

We consider a coded MU-MIMO system, which employs spatial multiplexing with $U$ UE and $B \geq U$ BS antennas. We assume that each of the $U$ UE has a single antenna. The information bits $\mathbf{b}$ of the $U$ UEs are encoded (e.g., using a convolutional code) and the resulting coded bit-stream $\mathbf{x}$ is mapped (using Gray labeling) to a sequence of transmit vectors $\mathbf{s} \in \mathcal{O}^U$, where $\mathcal{O}$ corresponds to the scalar complex constellation of size $2^Q$. Each transmit vector $\mathbf{s}$ is associated with $UQ$ binary values $x_{i,b} \in \{0,1\}$, $u = 1, \ldots, U$, $b = 1, \ldots, Q$, corresponding to the $b$th bit of the $u$th entry (i.e., spatial stream) of $\mathbf{s}$. We assume $\mathbb{E}_{\mathbf{s}}[\mathbf{s}\mathbf{s}^{\mathrm{H}}] = E_s \mathbf{I}_U$, where $E_s$ denotes the symbol variance. The baseband input-output relation of the wireless MIMO channel is given by $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}$ where $\mathbf{H}$ stands for the $B \times U$ complex-valued channel matrix, $\mathbf{y}$ is the $B$-dimensional received vector, and $\mathbf{n}$ is $B$-dimensional i.i.d. zero-mean complex Gaussian distributed noise with variance $N_0$ per entry. We assume that the channel matrix $\mathbf{H}$, the noise variance $N_0$, and the symbol variance $E_s$ are perfectly known at the receiver.[1]

### 7.2.2 Iterative MIMO Detection and Decoding

Iterative MIMO detection and decoding is based on the ideas of turbo-decoding [175]. Here, reliability information of the coded bits—in terms of

[1] In practice, channel-state information (CSI) is commonly acquired through training and hence, not perfect. Since imperfect CSI penalizes the performance of all considered MIMO detection algorithms in a similar way, we assume—for the sake of simplicity of exposition—perfect CSI throughout this chapter of the thesis.

log-likelihood ratios (LLRs)—is iteratively exchanged between the SISO detector and the SISO channel decoder to improve the error-rate performance. In each iteration, the SISO detector computes extrinsic LLRs [2, 4]

$$\Lambda_{i,b}^{\mathrm{e}} = \log\left(\frac{\mathrm{P}[x_{i,b} = 1 \mid \mathbf{y}]}{\mathrm{P}[x_{i,b} = 0 \mid \mathbf{y}]}\right) - \Lambda_{i,b}^{\mathrm{a}} \tag{7.1}$$

for the coded bits $x_{i,b}$, based on the received vector $\mathbf{y}$ and on the a-priori LLRs $\Lambda_{i,b}^{\mathrm{a}}$, $i = 1, \ldots, U$, $b = 1, \ldots, Q$. The extrinsic LLRs $\Lambda_{i,b}^{\mathrm{e}}$, which indicate the reliability for each coded bit $x_{i,b}$, are then delivered to the SISO channel decoder, which computes *new* a-priori LLRs $\Lambda_{i,b}^{\mathrm{a}}$, $\forall i, b$, that are used by the SISO detector in the next iteration. After a given number of iterations (denoted by $I$), the SISO channel decoder provides final decisions $\hat{\mathbf{b}}$ for the information bits $\mathbf{b}$ based on the LLRs at the channel-decoder output.[2]

## 7.3 Robust Algorithm Optimizations for LAMA

We will first present the LAMA algorithm and its limitations, and detail the algorithmic improvements for practical wireless channels.

### Posterior Mean and Variance Functions

The LAMA algorithm is first initialized by $\hat{\mathbf{s}}^1$ and $\hat{\gamma}_1^2$ which correspond to the mean and variance for the prior distribution $p(S)$ according to [1]: Therefore, for

[2] Early-termination schemes that reduce unnecessary iterations in turbo decoders (e.g., [176]) may also be used in iterative MIMO systems to reduce power consumption or to improve the (average) throughput.

each element $u = 1, \ldots, U$ we have the mean as:

$$\hat{s}_u^1 = \mathbb{E}_S[S] = \int_{\mathbb{C}} S p(S) \mathrm{d}S = \sum_{a \in \mathcal{O}} a \mathrm{P}[S = a] \,,$$

and the variance as:

$$\hat{\gamma}_1^2 = \mathbb{V}\mathrm{ar}_S[S] = \sum_{a \in \mathcal{O}} |a|^2 \mathrm{P}[S = a] - |\mathbb{E}_S[S]|^2 \,.$$

At every iteration of LAMA, we compute the posterior mean and variance function, $\mathsf{F}$ in (3.19) and $\mathsf{G}$ in (3.20), respectively, which operates element-wise on vector inputs. With the prior, the functions $\mathsf{F}$ and $\mathsf{G}$ are given by:

$$\begin{aligned} \mathsf{F}(\hat{z}, \hat{\gamma}^2) &= \sum_{a \in \mathcal{O}} a \mathrm{P}[S = a | \hat{z}] \,, \\ \mathsf{G}(\hat{z}, \hat{\gamma}^2) &= \sum_{a \in \mathcal{O}} |a|^2 \mathrm{P}[S = a | \hat{z}] - \left| \mathsf{F}(\hat{z}, \hat{\gamma}^2) \right|^2 , \end{aligned}$$

Here, the conditional probability $\mathrm{P}[S = a | \hat{z}]$ is computed as:

$$\mathrm{P}[S = a | \hat{z}] = \frac{\exp\left( -\frac{1}{\hat{\gamma}^2} |\hat{z} - a|^2 + \log \mathrm{P}[S = a] \right)}{\sum_{a' \in \mathcal{O}} \exp\left( -\frac{1}{\hat{\gamma}^2} |\hat{z} - a'|^2 + \log \mathrm{P}[S = a'] \right)}. \tag{7.2}$$

**Remark 1.** *$\mathsf{F}(z, 0)$ and $\mathsf{G}(z, 0)$ reduces to $\mathbb{E}_S[S]$ and $\mathbb{V}\mathrm{ar}_S[S]$.*

Remark 1 is beneficial as we can reuse the posterior mean and variance units for initialization of LAMA without having the need for a separate unit.

We note that for the $u$th user, the prior distribution $\mathrm{P}[s_u = a]$ is supplied by the decoder so that $\mathrm{P}[s_u = a] = \prod_{j=1}^{Q} \mathrm{P}\left[x_{u,j} = [a]_j\right]$, where $[a]_j \in \{0, 1\}$ denotes the $j$th bit of the symbol $a \in \mathcal{O}$. The detector provides the prior distribution in the form of

$$\mathrm{P}\left[x_{u,j} = [a]_j\right] = \frac{1}{2}\Big(1 + (2[a]_j - 1) \tanh(\Lambda_{u,j}^{\mathrm{a}}/2)\Big), \tag{7.3}$$

where the $\tanh(\Lambda/2)$ function can be efficiently implemented in hardware via look-up tables. Through numerical simulations, *intrinsic* a-priori LLRs for the

posterior mean and variance computations instead of the extrinsic ones to significantly better error-rate performance of LAMA; the same observation was found in [139]. Therefore, throughout the section, we will assume that the detector provides intrinsic a-priori LLRs.

#### Limitations

The LAMA algorithm presented Algorithm 2 requires each entry of the channel matrix $\mathbf{H}$ to be i.i.d. Gaussian with variance $1/B$. However, in practice each UE typically exhibits different large-scale fading (e.g., affected by the distance to the BS), resulting in channel matrices $\mathbf{H}$ whose columns have different scale. In addition, the large antenna array at the BS may exhibit correlation.

Finally, computing the posterior mean $\mathsf{F}$ and variance $\mathsf{G}$ functions require high dynamic range where even the use of double-precision floating-point arithmetic becomes numerically unstable in the high signal-to-noise regime.

### 7.3.1 Algorithmic Improvements For General Channels

#### Support For Non-Uniform Large-Scale Fading

We describe how LAMA is made robust to more general channel matrices. To do so, we will use ideas from the generalized AMP (GAMP) algorithm proposed in [59]. Unlike GAMP, which assumes arbitrary variances in the channel matrix, we only assume that each user experiences a difference variance (e.g. caused by large-scale fading). This assumption allows us to rewrite the channel matrix $\mathbf{H}$ as

$\widetilde{\mathbf{H}}\mathbf{D}$ where $\mathbf{H}$ is i.i.d. Gaussian and $\mathbf{D}$ is a diagonal matrix containing each user's large-scale fading.

**Damping For BS-side Correlation**

Despite the extension for non-uniform large-scale fading by GAMP, we perform another improvement to LAMA to enable robust performance for correlated channels. As suggested in [124,157,177], we perform damping of the intermediate variance parameters (see Algorithm 5 for $\theta$ parameters) to slow the updates. The damping parameters, $\theta_x$, $\theta_\rho$, and $\theta_\tau$ can be obtained through numerical simulations.

### 7.3.2 Efficient Posterior Mean and Variance Computation

As described in Chapter 3, the original LAMA function requires the computation of the posterior mean and variance, which require high numerical precision, which makes the design of high throughput architectures that operate on fixed-point arithmetic challenging . In this section, we detail the key optimizations that allow LAMA to be implemented economically in VLSI that deploy finite precision, but performs very-close to the LAMA in infinite precision.

**Reduction of Posterior Mean Sum: $|\mathcal{O}|$ to $2\sqrt{|\mathcal{O}|}$**

Without loss of generality, we will assume that the constellation is $M^2$-QAM with Gray coding.

We will first reduce the computational complexity by exploiting Gray coding where the data symbols are independent in the real and imaginary plane. Therefore, we use the separability property in Theorem 13 that enable us to represent the complex-valued posterior mean as two real-valued posterior means computed independently in real and imaginary domain, i.e.,

$$\mathsf{F}(\hat{z}, \hat{\gamma}^2) = \mathsf{F}^{\mathrm{R}}(\mathrm{Re}(\hat{z}), \hat{\gamma}^2/2) + \mathrm{i}\mathsf{F}^{\mathrm{R}}(\mathrm{Im}(\hat{z}), \hat{\gamma}^2/2) \tag{7.4}$$

The separable property has two advantages as it (i) reduces the sum over $M^2$ complex-valued entries to over two $M$ real-valued entries, and (ii) allows the computation to computed in parallel to reduce the critical path. For an example, for a 256-QAM constellation, (7.4) allows us to exactly compute the posterior mean by decomposing it as two posterior mean units for a 16-PAM constellation which can be computed concurrently.

For the case of binary phase-shift keying (BPSK), we use the separable property for QPSK as done in (7.4), and disregard the imaginary portion to get the desired result.

**Further Reduction By Bit-Domain Computation**

Although the complexity is reduced from $|\mathcal{O}|$ to $\sqrt{|\mathcal{O}|}$ from (7.4), computing the posterior mean function for the real-valued constellation $\mathcal{O}^{\mathrm{real}} = \mathrm{Re}(\mathcal{O})$ requires high numerical precision. This is due to the fact that the posterior distribution (7.2) requires a division of two functions that decay exponentially fast to 0 for large values of $\rho = 1/\hat{\gamma}^2$.

To reduce the precision requirement of the computation of (7.2), we use the decoupling property of LAMA and propose an approximation based on Gray

coding. We will derive our optimization for the real part; derivations for the imaginary follow exactly. To this end, we assume that each bit-wise entries is independent, i.e., for the $u$th user and any symbol $a^{\mathrm{R}} \in \mathcal{O}^{\mathrm{real}}$,

$$\mathrm{P}\left[\hat{s}_u^{\mathrm{R}} = a^{\mathrm{R}} | \hat{z}_u^{\mathrm{R}}\right] \simeq \prod_{j=1}^{Q/2} \mathrm{P}\left[x_{u,j}^{\mathrm{R}} = [a^{\mathrm{R}}]_j | \hat{z}_u^{\mathrm{R}}\right]. \tag{7.5}$$

Therefore, the bit-wise probabilities can be computed by:

$$\mathrm{P}\left[x_{u,j}^{\mathrm{R}} = b | \hat{z}_u^{\mathrm{R}}\right] = \sum_{a \in \mathcal{X}_j^b} \mathrm{P}\left[\hat{s}_u^{\mathrm{R}} = a^{\mathrm{R}} | \hat{z}_u^{\mathrm{R}}\right], \tag{7.6}$$

where $b \in \{0, 1\}$, and $\mathcal{X}_j^b$ contains all symbols $a^{\mathrm{R}} \in \mathcal{O}^{\mathrm{real}}$ such that $[a^{\mathrm{R}}]_j = b$, so that $\mathcal{X}_j^0 \cup \mathcal{X}_j^1 = \mathcal{O}^{\mathrm{real}}$.

**Computation of the A-Posteriori LLR through Max-Log Approximation**

Although (7.6) provides the expression for LLR computation for the decoder as in (7.1), this requires summation over $\mathcal{X}_j^1$ symbols, a division, and a $\log(\cdot)$ operation. Therefore, rather than computing $\mathrm{P}\left[x_{u,j} = 1 | \hat{z}_u\right]$ first, and then computing the a-posteriori LLR for (7.1), we propose a method to compute the LLR directly to reduce the complexity.

$$\begin{aligned}
\Lambda_{u,j}^{\mathrm{d}} &= \log \frac{\mathrm{P}\left[x_{u,j}^{\mathrm{R}} = 1 | \hat{z}_u\right]}{\mathrm{P}\left[x_{u,j}^{\mathrm{R}} = 0 | \hat{z}_u\right]} \\
&= \log \sum_{a \in \mathcal{X}_j^1} \mathrm{P}\left[\hat{s}_u^{\mathrm{R}} = a^{\mathrm{R}} | \hat{z}_u^{\mathrm{R}}\right] - \log \sum_{a \in \mathcal{X}_j^0} \mathrm{P}\left[\hat{s}_u^{\mathrm{R}} = a^{\mathrm{R}} | \hat{z}_u^{\mathrm{R}}\right].
\end{aligned}$$

As described in (7.3), $\mathrm{P}\left[x_{u,j}^{\mathrm{R}} = 1 | \hat{z}_u\right]$ can be computed efficiently from $\Lambda_{u,j}^{\mathrm{d}}$ via look-up tables. To compute $\Lambda_{u,j}^{\mathrm{d}}$ efficiently, we approximate $\Lambda_{u,j}^{\mathrm{d}}$ by using the

max-log approximation [139]:

$$\Lambda_{u,j}^{\mathrm{d}} \simeq \min_{a\in\mathcal{X}_j^1}\left\{\rho(\hat{z}_u^{\mathrm{R}}-a)^2 - \sum_{k=1}^{Q/2}\frac{2[a]_k-1}{2}\Lambda_{u,k}^{\mathrm{a}}\right\} - \min_{a\in\mathcal{X}_j^0}\left\{\rho(\hat{z}_u^{\mathrm{R}}-a)^2 - \sum_{k=1}^{Q/2}\frac{2[a]_k-1}{2}\Lambda_{u,k}^{\mathrm{a}}\right\}$$

In [139], the authors ignored the prior term and only considered $\Lambda_{u,j}^{\mathrm{d}} \simeq \min_{a\in\mathcal{X}_j^1}\rho(\hat{z}_u^{\mathrm{R}}-a)^2 - \min_{a\in\mathcal{X}_j^0}\rho(\hat{z}_u^{\mathrm{R}}-a)^2$. For LAMA, however, through numerical simulations, prior knowledge is crucial for robust error-rate performance. Thus, we make the following additional approximation[3]:

$$\Lambda_{u,j}^{\mathrm{d}} \simeq \rho\lambda_j^{\mathrm{post}}(\hat{z}_u^{\mathrm{R}}) + \lambda_{u,j}^{\mathrm{prior}},$$

where we introduced shorthand notation for $\lambda_j^{\mathrm{post}}(z)$ and $\lambda_j^{\mathrm{prior}}$,

$$\lambda_j^{\mathrm{post}}(z) = \min_{a\in\mathcal{X}_j^1}(z-a)^2 - \min_{a\in\mathcal{X}_j^0}(z-a)^2, \tag{7.7}$$

$$\lambda_{u,j}^{\mathrm{prior}} = \max_{a\in\mathcal{X}_j^1}\sum_{j=1}^{Q/2}\frac{2[a]_j-1}{2}\Lambda_{u,j}^{\mathrm{a}} - \max_{a\in\mathcal{X}_j^0}\sum_{j=1}^{Q/2}\frac{2[a]_j-1}{2}\Lambda_{u,j}^{\mathrm{a}}.$$

We note that the prior LLR $\lambda_{u,j}^{\mathrm{prior}}$ is given by the decoder so it is independent of the input $\hat{z}^{\mathrm{R}}$. The following theorem, with proof in Appendix A.5.1 shows that prior LLR $\lambda_{u,j}^{\mathrm{prior}}$ coincides exactly to the prior distribution $\Lambda_{u,j}^{\mathrm{a}}$.

**Theorem 43.** *The prior LLR expression $\lambda_{u,j}^{\mathrm{prior}}$ is equal to $\Lambda_{u,j}^{\mathrm{a}}$ provided by the channel decoder.*

Theorem 43 shows that LAMA with our optimization directly incorporates the prior $\Lambda_{u,j}^{\mathrm{a}}$ without additional processing. Thus, the LLR $\Lambda_{u,j}^{\mathrm{d}}$ for computation of posterior mean and variance is computed by:

$$\Lambda_{u,j}^{\mathrm{d}} = \rho\lambda_j^{\mathrm{post}}(\hat{z}_u^{\mathrm{R}}) + \Lambda_{u,j}^{\mathrm{a}}, \tag{7.8}$$

[3]The approximation yields no loss if we have equally-likely symbols.

where $\lambda_j^{\text{post}}(\hat{z}_u^{\text{R}})$ can be computed efficiently. We note that the extrinsic LLR for the decoder (7.1) is provided as $\Lambda_{u,j}^{\text{e}} = \Lambda^{\text{d}}$.

#### Computation of Posterior Mean and Variance

Now that we have detailed our key optimizations to transform the symbol-domain posterior distribution $\mathrm{P}\left[s_u = a^{\text{R}} | \hat{z}_u^{\text{R}}\right]$ and $\mathrm{P}\left[s_u = a^{\text{I}} | \hat{z}_u^{\text{I}}\right]$ to bit-domain LLRs, we finally summarize how we compute the posterior mean and variance functions.

Once the LLR terms $\Lambda_{u,j}^{\text{d}}$ for $j = 1, \ldots, Q/2$ for the real and $k = Q/2 + 1, \ldots, Q$ for the imaginary are obtained, we use look-up tables to compute the bit-wise probabilities $\mathrm{P}\left[x_{u,j}^{\text{R}} = [a^{\text{R}}]_j | \hat{z}_u^{\text{R}}\right]$. The final posterior mean is computed by:

$$\begin{aligned}\mathsf{F}(\hat{z}_u, \hat{\gamma}^2) = &\sum_{a^{\text{R}} \in \mathcal{O}^{\text{real}}} a^{\text{R}} \prod_{j=1}^{Q/2} \mathrm{P}\left[x_{u,j}^{\text{R}} = [a^{\text{R}}]_j | \hat{z}_u^{\text{R}}\right] \\ &+ 1i \sum_{a^{\text{I}} \in \mathcal{O}^{\text{imag}}} a^{\text{I}} \prod_{k=Q/2+1}^{Q} \mathrm{P}\left[x_{u,k}^{\text{I}} = [a^{\text{I}}]_k | \hat{z}_u^{\text{I}}\right],\end{aligned}$$

where the posterior variance function $\mathsf{G}(\hat{z}, \hat{\gamma}^2)$ can be computed similarly. These functions can be implemented in hardware at low complexity using the results [178].

### 7.3.3 Preprocessing

We perform preprocessing [178] on LAMA by reformulating the algorithm to directly run on the normalized Gram matrix $\widetilde{\mathbf{G}} = \mathbf{I}_U - \mathrm{diag}(\mathbf{G})^{-1}\mathbf{G}$ where the Gram matrix is $\mathbf{G} = \mathbf{H}^{\text{H}}\mathbf{H}$, large-scale fading vector $\mathbf{g} \in \mathbb{R}^U$ where $g_u = G_{u,u}/B$

for $u = 1, \ldots, U$, and the normalized matched filter $\widetilde{\mathbf{y}}^{\mathrm{MF}} = \mathrm{diag}(\mathbf{G})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y}$ without any performance loss. Preprocessing is beneficial to LAMA since in massive MU-MIMO, $U \ll B$, so LAMA operates on a lower-dimensional problem. Also, the precomputed normalized Gram matrix and the large-scale fading vector can be reused detection of other OFDM symbols and iterative detection and decoding. In addition, the resulting LAMA algorithm can be used for systems with various BS antenna configurations. The resulting algorithm is shown in Algorithm 5, where we modified the posterior functions $\mathsf{F}$ and $\mathsf{G}$ to incorporate a-priori LLRs $\mathbf{\Lambda}^{\mathrm{prior}}$, and we directly operate on the post-equalization SINR $\rho = \xi^{-2}$.

**Algorithm 5** (Robust LAMA)**.** *Initialize* $\tau^0, \rho^0, \tilde{\rho}^1 = 0$, $\hat{\mathbf{z}}^1 = \mathbf{0}_{U\times 1}$, $\theta_\tau^1, \theta_x^1, \theta_\rho^1 = 1$. *Then, for every iteration* $t = 1, 2, \ldots, t_{\max}$ *compute:*

$$
\begin{aligned}
\hat{\mathbf{s}}^t &= \mathsf{F}(\hat{\mathbf{z}}^t, \tilde{\rho}^t\mathbf{g}, \mathbf{\Lambda}^{\mathrm{a}}) \\
\hat{\mathbf{e}}^t &= \mathsf{G}(\hat{\mathbf{z}}^t, \tilde{\rho}^t\mathbf{g}, \mathbf{\Lambda}^{\mathrm{a}}) \\
\tau^t &= \theta_\tau^t \mathbf{g}^{\mathrm{T}}\hat{\mathbf{e}}^t + (1-\theta_\tau^t)\tau^{t-1} \\
\mathbf{a}^t &= \hat{\mathbf{z}}^t - \hat{\mathbf{s}}^t \\
\nu^t &= \tau^t \rho^{t-1} \\
\tilde{\mathbf{s}}^t &= \theta_x^t \hat{\mathbf{s}}^t + (1-\theta_x)\tilde{\mathbf{s}}^{t-1} \\
\tilde{\rho}^{t+1} &= \left[\theta_\rho^t \frac{1}{B}(\tau^t + N_0) + (1-\theta_\rho^t)\frac{1}{\tilde{\rho}^t}\right]^{-1} && (7.9) \\
\rho^{t+1} &= \left[\frac{1}{B}(\tau^t + N_0)\right]^{-1} && (7.10) \\
\hat{\mathbf{z}}^{t+1} &= \widetilde{\mathbf{y}}^{\mathrm{MF}} + \widetilde{\mathbf{G}}\tilde{\mathbf{s}}^t + \nu^t \mathbf{a}^t, && (7.11)
\end{aligned}
$$

*and* $(\theta_\tau^{t+1}, \theta_x^{t+1}, \theta_\rho^{t+1}) = (\theta_\tau, \theta_x, \theta_\rho)$. *The outputs are (i) decoupled output* $z_u^t$ *and (ii) post-equalization SINR* $\tilde{\rho}^t g_u$ *for each UE* $u = 1, \ldots, U$.

In this section, we describe a VLSI architecture for the LAMA algorithm and

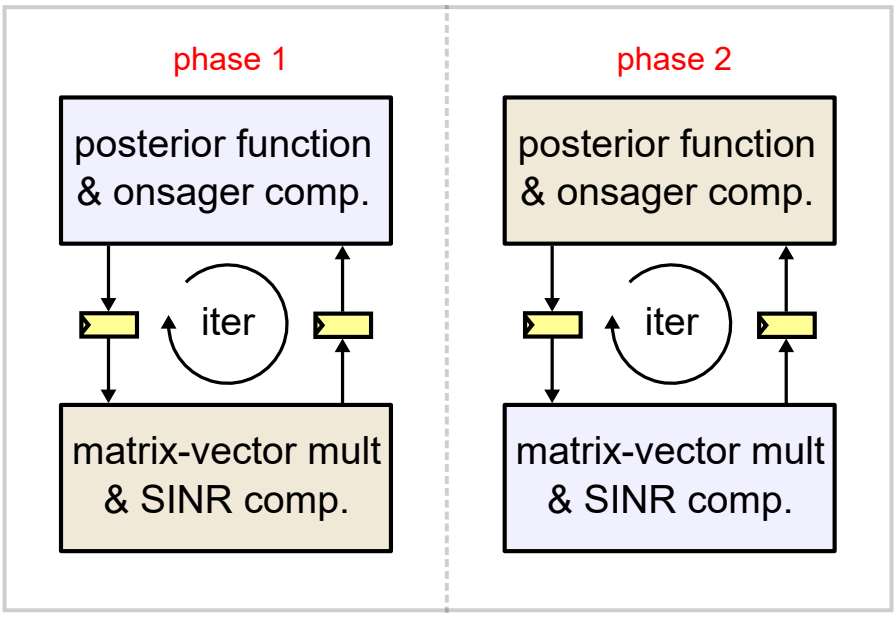



Figure 7.1: Coarse-grained pipeline architecture for LAMA: During the first phase, the posterior function unit and matrix-vector unit each solves a problem instance, and saves the outputs to the middle shift registers. During the second phase, the two units switch the problem instances. As a result, our design is able to achieve high-throughput, while keeping modularity.

present the key implementation details that enable us to achieve high throughput with low area in massive MU-MIMO systems.

### 7.3.4 Architecture Overview

#### Modular Design Methodology

In order to achieve high throughput and a modular design, we partition the LAMA into two units. Since the LAMA algorithm is iterative, each of the two units can execute units in a parallel and (course-grained) pipelined fashion. A graphical example of LAMA is shown in Fig. 7.1. For every iteration of LAMA, the processing is divided into two phases. During the first phase, the posterior

unit computes a problem instance (blue) and the matrix-vector multiplier solves another independent problem instance (brown). For both units, the outputs are stored in central shift-registers. During the second phase, both of the units take inputs from the central shift-registers, and thus, the posterior unit computes another problem instance (brown) and likewise for matrix-vector multiplier. Our design principle of LAMA is beneficial as it allows each unit to be designed, optimized, and verified independently, which ultimately results in a reduced development time.

**Top-level Architecture**

The detailed top-level architecture of Algorithm 5 is shown in Fig. 7.2. As discussed in Section 7.3.3, $U \ll B$ in massive MIMO so we perform preprocessing, which allows LAMA to run on a lower-dimensional problem without any performance loss. Such methods are beneficial as it reduces the storage and computational overhead without any performance loss. We note that the normalized Gram matrix operation can be done off-chip via highly decentralized architectures [149, 179].

The posterior and the matrix-vector multiplier unit are partitioned so that inputs are fed into the corresponding units only; the large-scale fading gain $\mathbf{g}$, a-priori LLRs $\mathbf{\Lambda}^{\mathrm{a}}$, and the noise variance $N_0$ are fed into the posterior unit, and the normalized Gram matrix $\widetilde{\mathbf{G}}$, matched filter $\widetilde{\mathbf{y}}^{\mathrm{MF}}$ are fed into the matrix-vector multiplier unit. The two units perform the assigned computations in $T_{\mathrm{s}}$ clock cycles and the results of each unit are exchanged for additional processing for subsequent iterations. For the last iteration, the outputs from the matrix-vector multiplier unit are sent to the LLR computation unit for LLR computation. The

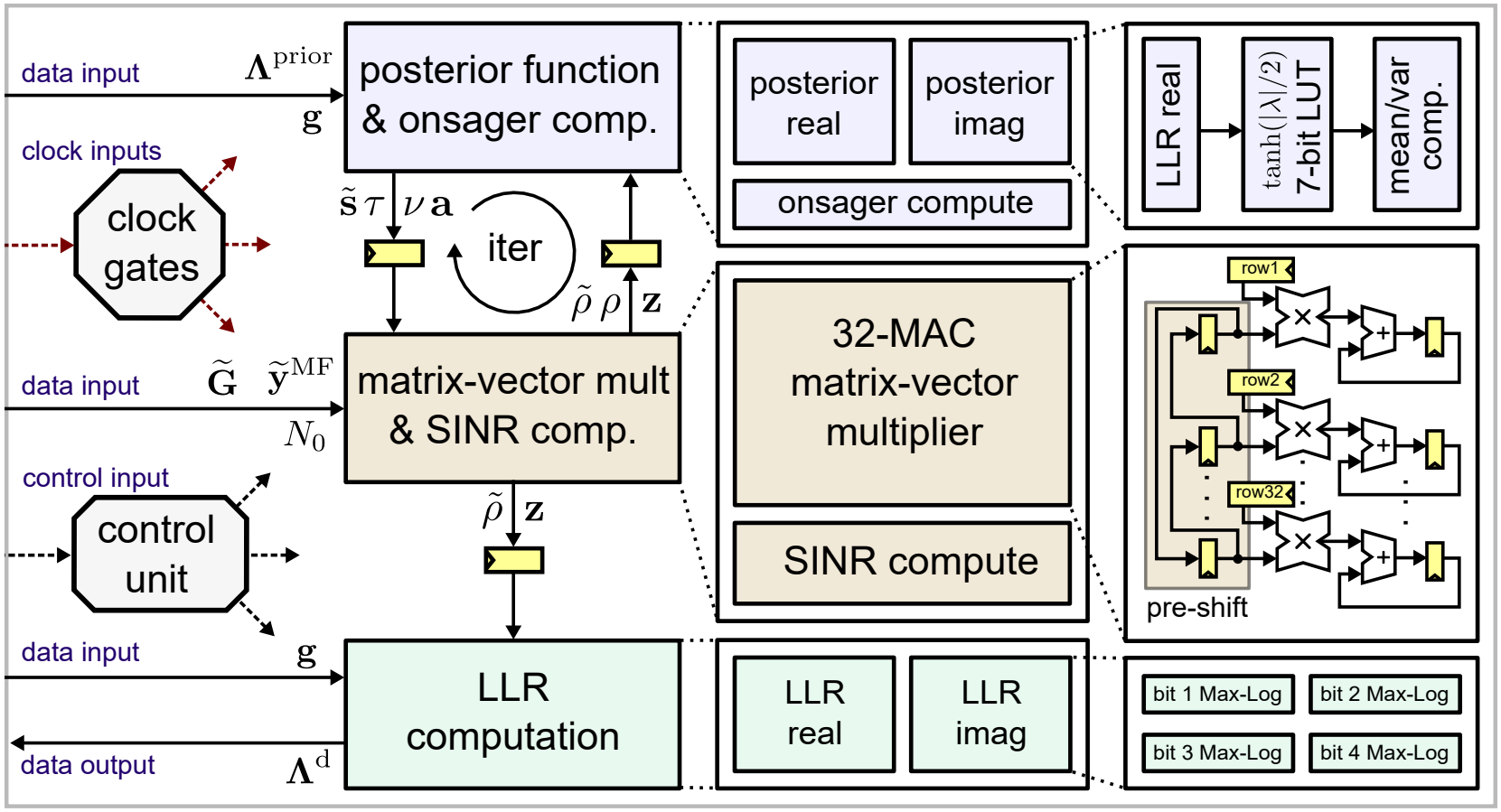



Figure 7.2: Top-level diagram of LAMA for Algorithm 5: Each iteration of LAMA is divided into two phases– posterior and matrix-vector multiplication. Once $t_{\max}$ iterations are reached, the outputs $\tilde{\rho}$ and $\hat{\mathbf{z}}$ are sent to the LLR computation unit. The computed LLRs $\mathbf{\Lambda}^{\text{d}}$ are then sent to the channel decoder. The posterior function, and LLR units operate separately on real and imaginary domain, which significantly reduces computational complexity, leading to a low-area and high-throughput design.

LLR computation takes $T_{\text{LLR}}$ clock cycles and thus, over the detector delivers a new set of $UQ$ LLR values resulting in a sustained throughput of

$$\Theta = \frac{UQ}{t_{\max} T_{\text{s}} + T_{\text{LLR}}} f_{\text{clk}} \quad \text{[bit/s]} \tag{7.12}$$

Since (7.12) scales linearly in the clock frequency $f_{\text{clk}}$ the throughput of the detector is maximized by minimizing the length of the critical path of the whole design. In order to arrive at low silicon area while maximizing the peak data-rate of data detection in 28nm CMOS technology, we chose $T_{\text{s}} = 36$ clock cycles.

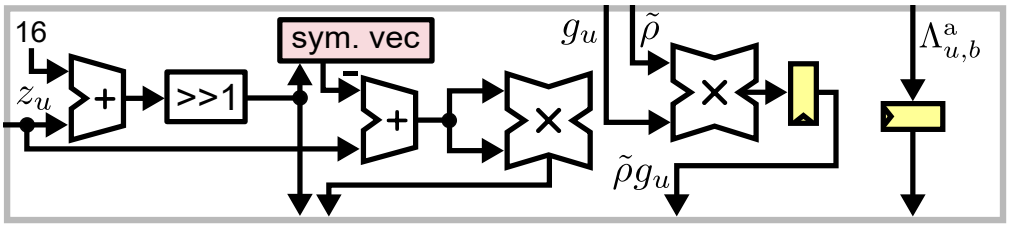
The control unit handles the input/output interface and damping parameter selection in the detector. To reduce dynamic power consumption in the case where no data-frame needs to be processed, the clock of each unit are gated individually.

**Input and Output Data Memories**

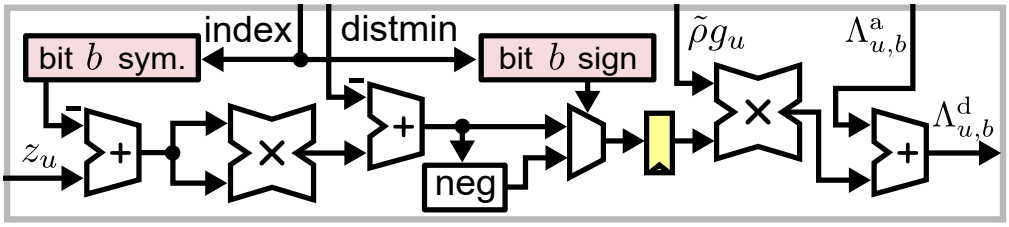
The input data memories for the LAMA algorithm store the necessary quantities for data detection, which include the precomputed Gram matrix $\widetilde{\mathbf{G}}$, matched-filter $\widetilde{\mathbf{y}}^{\text{MF}}$, large-scale fading gain $\mathbf{g}$ and the a-priori LLRs $\mathbf{\Lambda}^{\text{a}}$. The output data memories for LAMA store the a-posteriori LLRs $\mathbf{\Lambda}^{\text{d}}$, which will be sent to the decoder to perform channel coding and error correction. In order to minimize the circuit area, we use latch arrays built from standard cells [180].

**Intermediate Data Memories**

The data memories in the posterior and matrix-vector multiplier unit store all the intermediate values. In order to support a high memory bandwidth and




(a) Closest symbol index and post-equalization SINR calculation




(b) Max-log based LLR calculation for $b$th bit

Figure 7.3: Log-likelihood Ratio unit for $b$th bit in posterior function unit: (a) We compute the closest symbol index, the distance for a given QAM constellation, and the post-equalization SINR. (b) We use the precomputed quantities in (a) to compute the LLRs. LAMA directly incorporates prior the LLRs $\Lambda^{\mathrm{a}}$ for posterior function calculation.

to enable parallel access to multiple data words in an irregular manner, the memories are formed by arrays of flip-flops instead of using on-chip S-RAM macro-cells.

### 7.3.5 Posterior Function and Onsager Computation

We first divide the posterior function into real and imaginary parts as in (7.4) and compute them simultaneously. As discussed in Section 7.3.2, such decomposition reduces the computational complexity without any loss. The posterior function is divided into three stages, where we compute (1) LLR, (2) bit probabilities, and (3) posterior mean and variance, respectively.

**Bit-Domain Log-Likelihood Ratio Computation**

The architecture for the $b$th bit LLR compute unit is shown in Fig. 7.3. Since the largest constellation that LAMA supports is 256-QAM, corresponding to 16-PAM for real and imaginary separation, we have a total of 4 different LLR compute units corresponding to each bit for 16-PAM symbols in Gray coding. In order to build a compact implementation of (7.8), we employ max-log approximation to compute the posterior LLR $\lambda^{\text{post}}$. To compute (7.7) efficiently we use a two-step process where we (a) use numerical tricks to compute the closest symbol and its distance for the constellation used in data detection, and (b) use the symbol index to compute the distance for the complementary symbol.

Fig. 7.3(a) shows the architecture for the computation of the distance to the closest symbol. We first exploit that the input $z_u$ has dynamic range with 4 signed bits. We first add 16 and divide by 2, which effectively provides the index of the closest symbol $a^\star$, which is stored in a LUT in ascending order. The distance between $z_u$ is computed by a square unit $(z_u - a^\star)^2$ and broadcasted to the four LLR compute units. In addition, the post-equalization SINR $\tilde{\rho} g_u$ is computed and delayed by a cycle to synchronize with the LLR units.

Fig. 7.3(b) shows the remaining step for the LLR computation. The closest index of $a^\star$ goes into two LUT tables where the first provides the symbol for the complementary bit for $\bar{a}^\star$ for distance calculation. The second LUT provides the sign correction for the distance difference for $a^\star$ and $\bar{a}^\star$ prior to the final multiplication with the post-equalization SINR.

**LUT Table: LLR to Probabilities**

Once the LLR for each bit values are computed, the LLRs are sent to a LUT to generate the bit-wise probabilities $\mathrm{P}\left[x_{u,j}^{\mathrm{R}}=1\right]$ for each bit $j=1,\ldots,4$. Instead of computing $\mathrm{P}\left[x_{u,j}^{\mathrm{R}}=1\right]$ directly from relation (7.6), we use the results form [178] to operate on $1-2\mathrm{P}\left[x_{u,j}^{\mathrm{R}}=1\right]$ for posterior mean and variance computation. By doing so, we only need to compute $\pm\tanh(\Lambda_{u,j}^{\mathrm{a}}/2)$, which can be computed easily using LUT. We utilize symmetry in the $\tanh(\cdot)$ function note that $\pm\tanh(\Lambda_{u,j}^{\mathrm{a}}/2)=\mathrm{sign}(\Lambda_{u,j}^{\mathrm{a}})\tanh(|\Lambda_{u,j}^{\mathrm{a}}|/2)$. Thus, the probabilities are generated from a 7-bit LUT table that contains the values $\tanh(x/2)$ for positive $x$. The computed bit-wise probabilities are then sent the mean and variance computation unit (see top right at Fig. 7.2).

**Bit-Wise Probabilities to Posterior Mean and Variance**

To compute the final posterior mean and variance, we use the approach proposed in Fig. 7.2, where we decompose the mean $\mathsf{F}$ and variance $\mathsf{G}$ functions into product and addition of the bit-wise probabilities. To support various constellations ranging from BPSK to 256-QAM, we implement the chip for 256-QAM and multiplex the intermediate signals to support smaller PAM constellations. To do so in a efficient manner, on top of the bit-wise probabilities $b_j^{-}=1-2\mathrm{P}\left[x_{u,j}^{\mathrm{R}}=1\right]$ for $j=1,2,3,4$, we introduce intermediate quantities

$$
\begin{aligned}
b_k^{+} &= 1+2\mathrm{P}\left[x_{u,k}^{\mathrm{R}}=1\right]=2-b_k^{-}\\
b_k^{\mathrm{byp}} &= \mathrm{EN}_k^{\mathrm{byp}}+(1-\mathrm{EN}_k^{\mathrm{byp}})b_k^{+}\\
C_1^{\mathrm{byp}} &= b_1^{\mathrm{byp}}\gg K_1^{\mathrm{byp}},
\end{aligned}
$$

for $k = 1$ and 3, so that the posterior mean function can be factored efficiently as:

$$\mathsf{F}(\mathbf{x}_{\mathbf{u}}^{\mathrm{R}}, \rho) = b_0^- (C_1^{\mathrm{byp}} + (\mathrm{EN}_1 b_1^-)(\mathrm{EN}_2 b_2^-) b_3^{\mathrm{byp}}).$$

The posterior variance function can be computed similarly. The choice of the enable signals EN and the shift constant $K_1^{\mathrm{byp}}$ enable the computation of $\mathsf{F}$ for BPSK, and $M$-PAM constellations for $M = 4, 8, 16$. We reiterate that our implementation of the posterior mean and variance function yields no approximation loss for BPSK/QPSK constellation.

### 7.3.6 Matrix-vector Multiplier Unit

As shown in (7.11), the matrix-vector multiplier unit computes $\mathbf{z}^{t+1} = \widetilde{\mathbf{y}}^{\mathrm{MF}} + \widetilde{\mathbf{G}}\tilde{\mathbf{s}}^t + \nu^t \mathbf{a}^t$ at every iteration of LAMA. The $32 \times 32$ complex-valued multiplication is carried out using 32 complex-valued multiply-accumulate (MAC) units; the matrix-vector operation is carried out on a column-by-column basis so that each MAC unit is associated to a row of the Gram matrix.

A straightforward approach to compute the $32 \times 32$ matrix-vector product for $\widetilde{\mathbf{G}}\tilde{\mathbf{s}}$ would be to broadcast each $j$th entry $\tilde{s}_j$ to all 32 individual MAC units for $j = 1, \ldots, 32$, and then perform column-wise MAC operation, completing the matrix-vector product over 32 cycles. Although this approach yields the correct solution, it may not be desirable for high-throughput as the input shift register $\tilde{s}_j$ has to be fed into all 32 MAC units, increasing fan-in. In addition, this approach requires the 32 element $\tilde{\mathbf{s}}$ to be stored over 32 cycles. Once the $32 \times 32$ matrix-vector product $\widetilde{\mathbf{G}}\tilde{\mathbf{s}}$ is computed, we utilize the same unit to compute $\nu^t \mathbf{a}^t$.

**Cannon's Algorithm**

To reduce the fan-in and the storage requirement of $\tilde{\mathbf{s}}$, we utilize a variant of Cannon's algorithm [42]. To this end, we store the $32 \times 32$ matrix $\widetilde{\mathbf{G}}'$ where each row is shifted by its index. To compute $\widetilde{\mathbf{G}}'\tilde{\mathbf{s}}$, the input $\tilde{\mathbf{s}}$ is first loaded into the input shift registers (the pre-shift block). We then circularly shift the entries of this input shift register (see illustration in Fig. 7.2) while sequentially calculating the MAC operations with the entries of matrix $\widetilde{\mathbf{G}}'$; the outputs are accumulated in the registers at the output of each MAC unit. This effectively implements a column-by-column matrix-vector operation in 32 clock cycles. In addition, our approach reduces fan-in, as each input shift register is connected to each MAC unit, and removes the need to store $\tilde{\mathbf{s}}$; all the entries of $\tilde{\mathbf{s}}$ are stored in the input shift registers.

### 7.3.7 SINR Computation: Newton-Raphson Based Reciprocal

In each iteration of LAMA, division operations are required for computation of $\tilde{\rho}$ and $\rho$ in (7.9) and (7.10) in Algorithm 5. We note that division operations are, in general, not well-suited for fixed-point implementation and off-the-shelf division circuits usually entail a large area, high latency, and/or long critical path. Since our main objective is to maximize the clock frequency of the LAMA data detector, we follow [139] and build a custom division unit that is able to compute reciprocal $1/\tau$ at a high throughput with a precision that results in a small implementation loss.

### Algorithm

Before we delve into the reciprocal algorithm, we first examine the structure of $\tilde{\rho}$ and $\rho$ that allows us to simplify the division operation. First, we note that $\tilde{\rho} = \rho$ if $\theta_\rho = 1$, hence we formulate our algorithm for $\tilde{\rho}$ only. A straightforward implementation of $\tilde{\rho}^{t+1}$ for LAMA would require two divisions which require, in general, high numerical precision. We reduce the divisions to one with low numerical precision first factoring out $\tilde{\rho}^t$ so that $\tilde{\rho}^{t+1} = \tilde{\rho}^t \left[\theta_\rho^t \tilde{\rho}^t \frac{1}{B}(\tau^t + N_0) + (1 - \theta_\rho^t)\right]^{-1}$. Since $b^t = \tilde{\rho}^t \frac{1}{B}(\tau^t + N_0) \in [0, 1]$ by construction, we first compute $b^t$ and then invert $\theta_\rho^t b + (1 - \theta_\rho) \in [1 - \theta_\rho, 1]$.

We implement the reciprocal unit by first generating an initial guess for $1/\tau$ from a look-up table (LUT) and a single iteration of Newton-Raphson iteration [139, 181]. We first shift the input value $\tau$ according to $\tau' = 2^\alpha \tau$, $\alpha \in \mathbb{Z}$ so that $\tau' \in [0.5, 1)$. This shifting operation results in $1/\tau' \in (1, 2]$ and thus, the subsequent computations can be carried out with high numerical stability[4]. Based on initial guess $\tau_0'$ of $1/\tau'$ from the LUT, we perform a single iteration of Newton-Raphson so that $\hat{\tau}' = 2\tau_0 - \tau' \tau_0^2$; the final result $\hat{\tau}'$ corresponds to an approximation of $1/\tau'$

### Architectures

In our design, the two divisions, $\tilde{\rho}$ and $\rho$, need to be completed over 36 clock cycles. Our numerical simulations have shown that 14-bit precision (excluding the initial shift) is sufficient to arrive at a negligible implementation loss. In addition, the initial guess is generated from a 5-bit LUT. Since we are only

[4] We multiply the output $1/\tau'$ by $2^\alpha$ with the aid of arithmetic shifters to get the final output $1/\tau$.

performing a single iteration of Newton-Raphson, we implement a pipelined architecture of the division unit, where a divison consumes 3 clock cycles.

### 7.3.8 LLR Computation Unit

Once the final iteration of LAMA is completed, the decoupled output $\mathbf{z}$ and the post-equalization SINR $\rho^t$ is sent to the LLR computation unit (see bottom green box in Fig. 7.2). The LLR computation unit for LAMA's output is identical to that in the posterior mean and variance, except the prior LLR is not used for the output LLRs. The final extrinsic LLRs for decoder provided in (7.1) is thus equivalent to a-posteriori LLRs $\Lambda^{\mathrm{d}}$ computed by LAMA. We include a separate LLR unit, rather than re-using the LLR compute unit in the posterior mean and variance unit, to maintain high sustained throughput shown in (7.12).

## 7.4 Implementation Results

The LAMA algorithm, along with all improvements and optimizations detailed in the previous sections, was fabricated in 28 nm CMOS technology. Fig. 7.4 details the post-placement layout with the highlighted processing units (PUs). The post-placement layout of the ASIC assumes preprocessing of the Gram matrix, matched filter, and large-scale fading gain to support SISO detection of 32 spatial streams with BPSK, QPSK, 16-, 64-, and 256-QAM modulation.

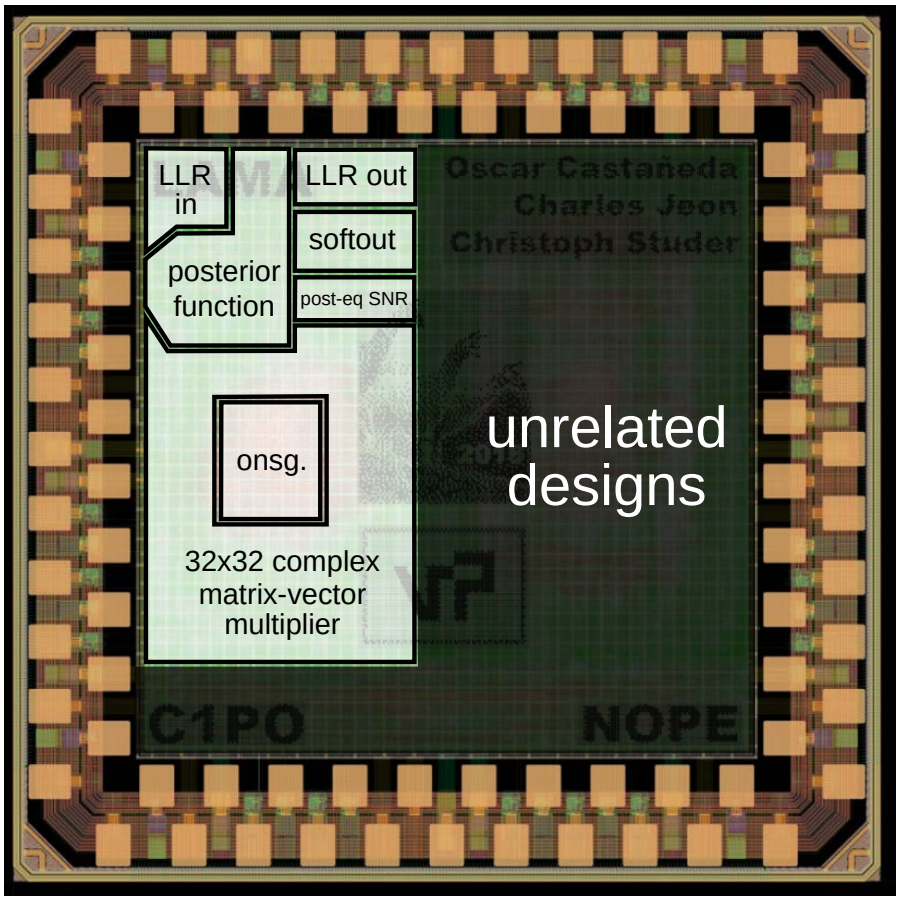



Figure 7.4: Post-placement layout of the LAMA chip with highlighted processing units

### 7.4.1 Post-placement Layout Results

The post-placement layout of the ASIC has the following key characteristics (see the detailed comparisons in Table 7.2). The core area is 0.37 $mm^2$ corresponding to 707 kGE. A detailed breakdown of each processing unit is shown in Table 7.1. The $32 \times 32$ complex-valued matrix-vector multiplier requires 15.2% and 26.3% of the silicon area with and without the I/O logic, respectively. We note that a significant portion of LAMA is dedicated to latched-based memory cells. The input and output LLRs, as well as Gram-matrix consume 27.8% and 48.2% of the silicon area with and without I/O logic, respectively.

Table 7.1: Detailed area and power breakdown of the post-placement layout of LAMA

| Processing unit | kGE | % | mW | % |
|---|---|---|---|---|
| Posterior function | 72.4 | 5.9 | 32.2 | 14.0 |
| Variance & Onsager computation | 62.5 | 5.1 | 14.3 | 6.2 |
| LLR input mem | 29.6 | 2.4 | 0.5 | 0.2 |
| Large-scale fading mem | 1.9 | 0.2 | 0.1 | 0.1 |
| Matrix-vector multiplier | 186.0 | 15.2 | 165.6 | 72.2 |
| Gram-matrix memory | 280.1 | 22.9 | 0.3 | 0.1 |
| SINR computation | 9.6 | 0.8 | 0.3 | 0.1 |
| LLR computation | 31.0 | 2.5 | 11.7 | 5.1 |
| LLR output mem | 29.5 | 2.4 | 4.1 | 1.8 |
| Miscellaneous[a] | 520.0 | 42.5 | 0.4 | 0.2 |
| Total | 1222.7 | 100 | 229.6 | 100 |

[a]Denotes logic used for the input/output-interface of the chip.

### Throughput

The maximum (internal) clock frequency of the post-placement layout of LAMA is 500 MHz, which results in a peak throughput of 443 MHz, which corresponds to the case of $256 \times 32$ massive MU with 256-QAM constellation at 8 LAMA iterations.[5]

### Power consumption

The power consumption[6] of LAMA according to each processing unit is shown in Table 7.1. The $32 \times 32$ complex-valued matrix-vector multiplier accounts for

[5]We note that higher throughput is achieved if we run less number of LAMA iterations. Depending on the operating SNR, this reduction of LAMA iterations will not impact packet error performance. However, since the operating SNR is not given, in general, we performed numerical simulations to obtain the minimum number of LAMA iterations where the packet error rate performance does not improve.

[6]Post-placement simulated at $V_{\text{dd}} = 0.9$ V core supply of the post-placement layout of LAMA.

72.2% of the total power consumption at 165.6 mW. Similarly, the posterior function also accounts for 14.0% of the total power consumption at 32.2 mW. These power consumption estimation results are to be expected, as these processing units perform the most complicated tasks in the data detector.

### 7.4.2 Fixed-point Error Rate Performance

In order to achieve near-optimal error-rate performance with fixed-point arithmetic, we simulated the LAMA algorithm to minimize the word-lengths. Since word-lengths of each variable in the LAMA algorithm is highly dependent on the channel model and massive MU-MIMO system parameters, we performed optimizations for $32 \times 32$ system for 16-QAM with typical urban micro scenario in WINNER II channel model [182]. Although majority of the existing literature and their corresponding implementations assume i.i.d. Rayleigh fading for the massive MIMO wireless channel, these assumptions are not practical, nor evident in real-world measurements [55].

The key optimized parameters are follows: We use 11 bits for the input and output LLRs. The real and imaginary parts of the precomputed normalized Gram matrix $\widetilde{\mathbf{G}}$, normalized matched filter $\widetilde{\mathbf{y}}^{\mathrm{MF}}$ are 14 and 16 bits respectively. The large-scale fading gain vector $\mathbf{g}$ has a precision of 9 bits.

#### Error Rate Performance

We now compare the packet error rate (PER) versus signal-to-noise performance of iterative MIMO decoding using the ideal (floating-point algorithm) Algo-

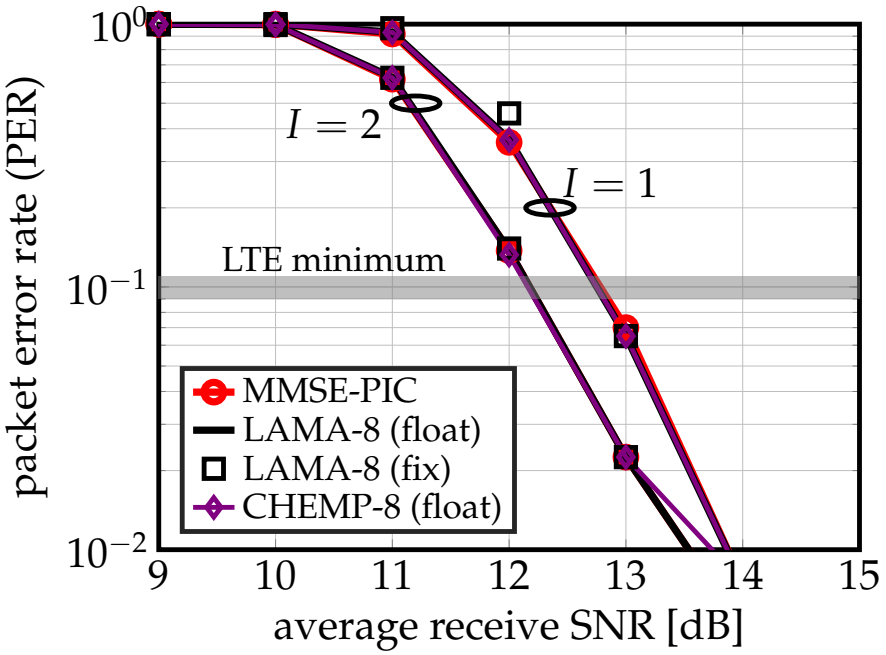



Figure 7.5: $256 \times 32$ MU-MIMO-OFDM system with i.i.d. Rayleigh fading and $R = 0.5$ 256-QAM constellation. Both floating and fixed-point implementation of LAMA and CHEMP achieve identical packet error rate performance as that given by MMSE-PIC.

rithm 5 and the corresponding (fixed-point) post-placement layout implementation. To compare the error-rate performance of our LAMA detector, we compare our results to MMSE-PIC [139], as well as CHEMP [111] that has been implemented into an ASIC in [35, 36] to showcase how existing massive MU-MIMO data detectors perform in real-world channel environments.

Figures 7.5 and 7.6 show the packet-error performance for a $256 \times 32$ massive MU-MIMO system with $R = 0.5$ 256-QAM constellation for i.i.d. Rayleigh fading and a more WINNER II [182] channel model. We simulate a LTE-like system, where we assume an OFDM system with active 1200 subcarriers and 2 OFDM symbols per packet, totaling 9600 data bits per packet. We use a rate $R = 0.50$ convolutional code for encoding and max-log BCJR algorithm for decoding [183]. For Fig. 7.6, we assume a typical urban micro scenario for the WINNER II [182]

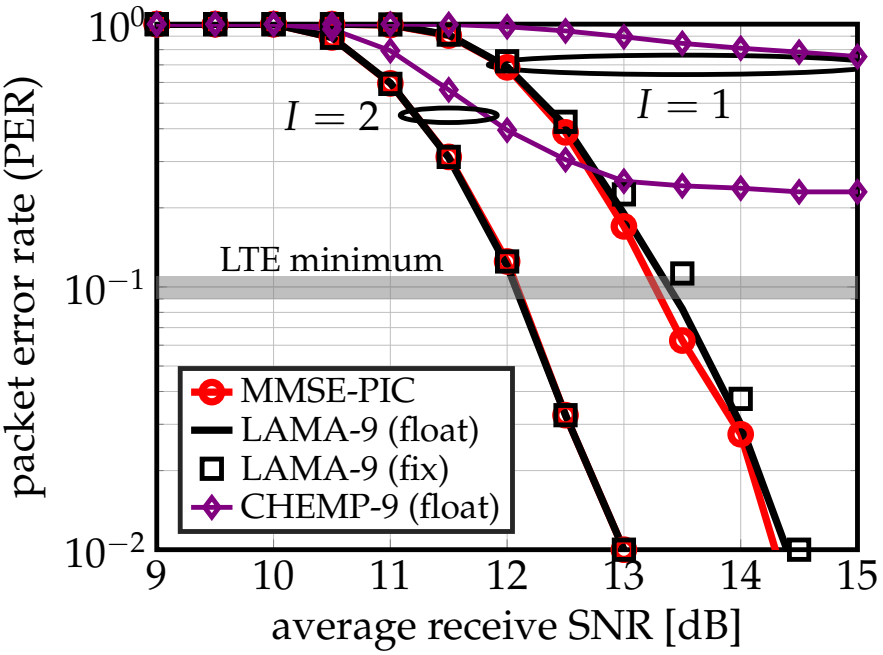



Figure 7.6: $256 \times 32$ MU-MIMO-OFDM system with WINNER II Urban Micro channel model and $R = 0.5$ 256-QAM constellation. Fixed point implementation of LAMA performs identically as floating implementation of LAMA. Similar to Fig. 7.8, CHEMP [111] results in error floors above LTE minimum even with one iteration of MIMO decoding.

channel model with BS antenna spacing of 10 m.

For the i.i.d. Rayleigh fading scenario shown in Fig. 7.5, we note that both floating and fixed-point implementations of LAMA, as well as CHEMP, achieve identical packet-error rate performance as that given by MMSE-PIC. In addition, we note a single iteration of iterative MIMO decoding provides an SNR gain of less than 1 dB.

For the typical urban micro scenario shown in Fig. 7.6, however, we first note that CHEMP results in an error floor above the LTE minimum of 10% even with iterative MIMO decoding. Both implementations of LAMA achieve identical packet-error rate performance as that given by MMSE-PIC. This striking difference between CHEMP and LAMA results from using a more realistic

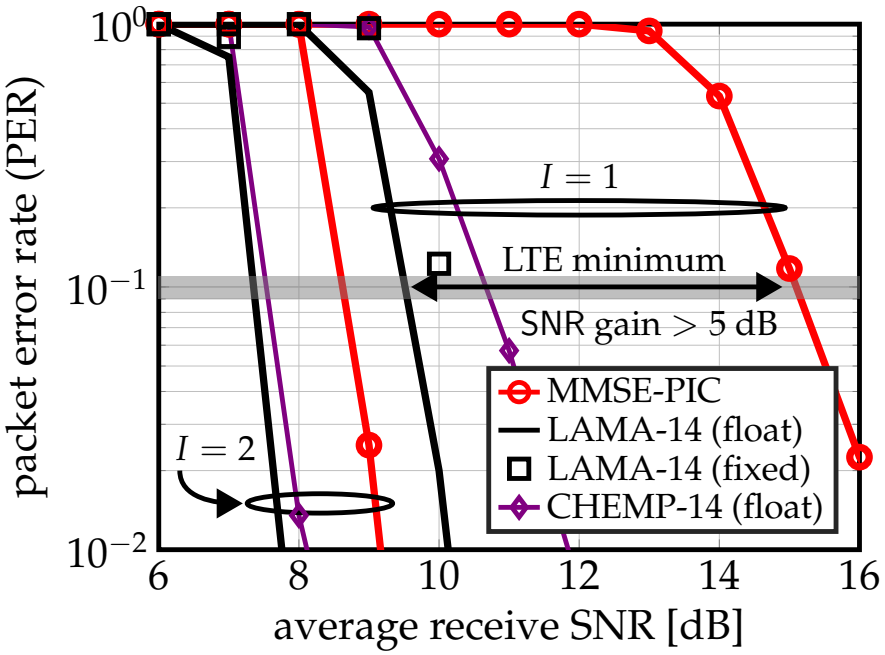



Figure 7.7: $32 \times 32$ MU-MIMO-OFDM system with i.i.d. Rayleigh fading and $R = 0.75$ QPSK constellation. Fixed point implementation of LAMA performs within 0.2 dB of floating implementation of LAMA, and performs 0.1 dB compared to CHEMP.

channel model, where the effects of shadow-fading, BS-side correlation are present. We note that for this $256 \times 32$ massive MU-MIMO system, LAMA is able to achieve error-rate performance of MMSE-PIC without having to do an matrix inversion.

Similarly, figures 7.7 and 7.8 show the packet-error performance for $32 \times 32$ massive MU-MIMO system with $R = 0.75$ QPSK constellation for both channel models with 3600 bits per packet. As identical to that given by previous experiment, we assume a typical urban micro scenario for the WINNER II [182] channel model with BS antenna spacing of 10 m.

For the i.i.d. Rayleigh fading scenario shown in Fig. 7.7, LAMA and CHEMP achieve the same LTE minimum packet error rate at 5 dB, and 2 dB lower SNR for

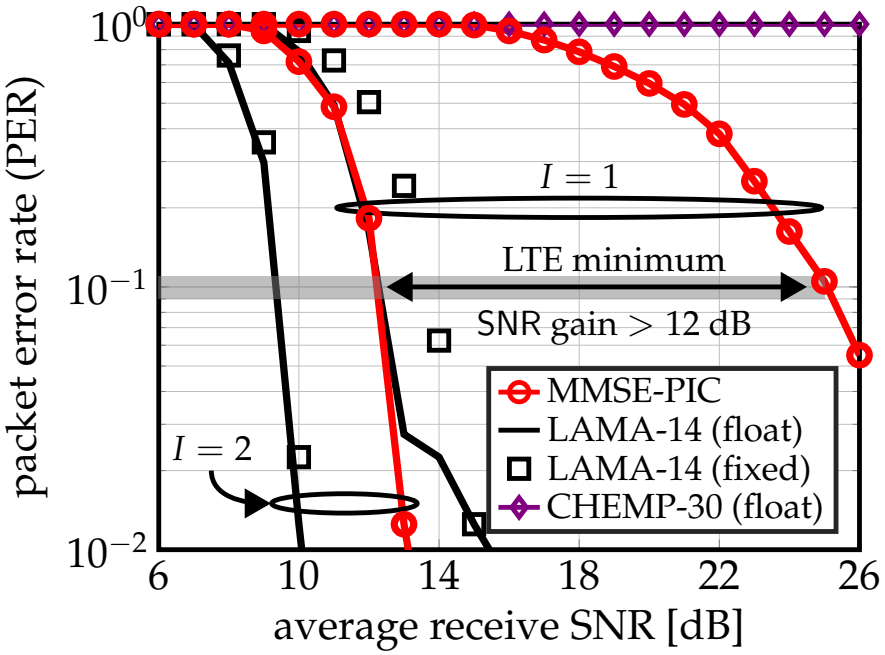



Figure 7.8: $32 \times 32$ MU-MIMO-OFDM system with WINNER II Urban Micro channel model and $R = 0.75$ QPSK constellation. Fixed point implementation of LAMA performs within 2 dB of floating implementation of LAMA. CHEMP [111] results in error floors even with one iteration of MIMO decoding.

1 and 2 iterations of MIMO decoding, respectively. In addition, the fixed point implementation of LAMA performs within 0.2 dB of floating implementation.

For the typical urban micro scenario shown in Fig. 7.8, however, CHEMP results in an error floor of 1 regardless of SNR and 30 iterations, whereas our LAMA implementations provide more than 12 dB gain in SNR at 14 iterations. In addition, the fixed point implementation of LAMA performs within 2 dB of floating implementation. Therefore, for this square $32 \times 32$ massive MU-MIMO system, LAMA is able to achieve more than 5 dB SNR gains in packet-error rate compared to CHEMP and MMSE-PIC in both channel models, without having to do an matrix inversion.

### 7.4.3 Comparisons to Other Detector Implementations

#### Throughput

For throughput figures reported in Table 7.2, we note LAMA achieves lower throughput than that given by another message-passing based method, CHEMP. This difference results from the fact that the authors in [35] have simplified their algorithm by channel hardening, approximating the posterior mean updates, and not supporting soft-input and soft-outputs. The authors in [36] support soft input and outputs, but does so only for QPSK constellation. The CHEMP algorithm strongly relies on the channel hardening assumption to provide an improved error-rate performance over linear data detectors. As demonstrated in Fig. 7.6 and Fig. 7.8, CHEMP, and thus, the approximated algorithms shown in [35] and [36], resulted in error floors above the LTE minimum for reliable communication.

Compared to linear methods [32, 34] and semidefinite-relaxation based method [173], LAMA achieves higher throughput when adjusted for the number of UEs. The main computational complexity for linear methods come from the inversion of the Gram matrix. Moreover, it is important to note that our proposed implementation, contrast to methods in [32, 34, 173], supports iterative MIMO decoding, which can provide significantly improved error-rate performance.

Finally, we compare our proposed LAMA implementation to that of [37], which utilizes another message-passing based method, expectation-propagation (EPD) [174]. Although both algorithms, EPD and LAMA supports real-world channels, the primary difference between EPD and LAMA is that EPD requires the computation of the matrix inverse every iteration. Since the complexity

of matrix inversion scales cubicly in the number of UEs [171], it will quickly dominate the data detection complexity in large number of UEs. Therefore, if we adjust the difference in the number of UEs cubicly, LAMA still achieves 2× greater throughput.

**Area**

For the area results reported in Table 7.2, it is important to note that the proposed LAMA implementation is the only detector capable of supporting a 32-UE soft input and output capability. When accounted for technology scaling, LAMA exhibits slightly greater area than that of the 32-UE implementation of CHEMP [35], but it is important to note that we support soft-input and soft-output capabilities and does not result in error-flooring behavior in practical systems. Compared to linear data detectors, LAMA exhibits significantly lower area as LAMA does not need to compute matrix inversion. It is important to note that the linear detectors achieve high area for a 8 UE system, whereas LAMA supports 32 UEs. Also, the EPD implementation exhibits 5× greater area for a 16 UE system due to the computation of the matrix inversion explicitly. Therefore, LAMA performs competitively in terms of area for data detection of a massive MU-MIMO system with 32 UEs.

**Area-Efficiency**

In terms of normalized area-efficiency, where we both scale the technology and number of UEs, LAMA performs better than linear, SDR and EPD methods. We achieve 17× and 6× greater area-efficiency compared to SDR and EPD,

respectively. We note that LAMA will exhibit higher area-efficiency if we account for cubic-scaling, which is the complexity of Gram matrix inversion. Although LAMA exhibits lower area-efficiency than that given by CHEMP-based methods, we note that these methods do not work on practical channels, as verified in our experiments. Therefore, we conclude that our design is highly competitive to other data detection implementations for massive MU-MIMO (which are unable to support iterative MIMO decoding).

## 7.5 Conclusions

In this chapter of the thesis, we presented a novel low-complexity soft-input soft-output (SISO) approximate message passing-based algorithm for massive MU-MIMO systems. We extended our existing asymptotically-optimal LAMA algorithm [38,39] to mitigate performance losses in finite dimensional and real-world channel models. In order to ensure high-throughput VLSI design at low-area employing fixed-point arithmetic, we detailed the key approximations and optimizations. We described the corresponding VLSI architecture and the post-placement layout results of LAMA in 28 nm CMOS technology. We demonstrate the efficacy of our algorithm by comparing real-world costs (in terms of silicon area and power) to other existing state-of-the-art massive MU-MIMO detectors. The proposed LAMA algorithm was shown to enable significant gains in spectral efficiency over recently proposed implementations, especially in real-world channel models. The uplink simulations proposed in this chapter demonstrates that near-optimal data detection is challenging, but feasible in realistic channel models with a large number of UEs.

Table 7.2: Post-layout results of LAMA and comparison to existing massive MU-MIMO data detectors

| | This work | Tang [35] | Chen [36] | Prabhu [34] | Peng [32] | Castañeda [173] | Tang [37] |
|---|---|---|---|---|---|---|---|
| System ($B \times U$) | $256 \times 32$ | $128 \times 32$ | $128 \times 8$ | $128 \times 8$ | $128 \times 8$ | $128 \times 8$ | $128 \times 16$ |
| Data Detection Alg. | LAMA | CHEMP | CHEMP | MMSE/ZF[a] | MMSE | SDR | EPD[b] |
| Iterative MIMO dec. | yes | no | yes | no | no | no | no |
| Modulation | 256-QAM | 256-QAM | QPSK | 256-QAM | 64-QAM | QPSK | 256-QAM |
| Preproc. quantities | Gram mat. | Gram mat. | Gram mat. | – | – | Gram mat. | Gram mat. |
| Supports real-world channels | yes | no | no | yes | yes | yes | yes |
| Provably optimal | yes | no | no | no | no | no | no |
| CMOS Tech. [nm] | 28 | 40 | 40 | 28 | 65 | 45 | 28 |
| Result | post-layout | silicon | silicon | silicon | silicon | post-layout | silicon |
| Core area [$mm^2$] | 0.36 | 0.58 | 0.076 | 1.10 | 2.57 | 0.48 | 2.0 |
| Frequency [MHz] | 500 | 425 | 500 | 300 | 680 | 560 | 569 |
| Power [mW] | 229.6 | 220.6 | 77.89 | 18 | 650 | 87.10 | 127 |
| Throughput [Gb/s] | 0.44 | 2.76 | 8 | 0.30 | 1.02 | 0.13 | 1.80 |
| Area Eff. [Gb/s/$mm^2$] | 1.21 | 4.76 | 105.26 | 0.27 | 0.40 | 0.26 | 0.90 |
| Norm. Area Eff.[c] [Gb/s/$mm^2$] | 1.21 | 13.87 | 19.18 | 0.02 | 0.31 | 0.07 | 0.23 |

[a]also supports precoding; [b]expectation-propagation; [c]Technology normalized to 28 nm CMOS technology assuming $f_{\text{clk}} \sim s$, $A \simeq 1/s^2$, and $P_{\text{dyn}} \sim (1/s)(V_{\text{dd}}/V'_{\text{dd}})$ and $U^2/32^2$

CHAPTER 8

# ASIC DESIGN OF A NONPARAMETRIC EQUALIZER FOR MASSIVE MU-MIMO

## 8.1 Introduction

It is widely believed that massive multi-user multiple-input multiple-output (MU-MIMO) will be a core technology for fifth-generation (5G) wireless systems. Massive MU-MIMO relies on base-station (BS) architectures with hundreds of antenna elements and radio-frequency (RF) chains that serve tens of user equipments (UEs) in the same time-frequency resource. While this emerging technology enables unprecedented spectral efficiency by means of fine-grained beamforming [5,30], it also poses significant practical implementation challenges.

### 8.1.1 The Case for Nonparametric Equalization

Data detection in the uplink (UEs transmit data to the BS) is among the most critical tasks from a spectral efficiency and hardware complexity perspective [22]. While optimal MIMO data detection is known to be NP-hard [58], it has been shown in [6,22] that linear minimum mean-square error (L-MMSE) equalization enables near-optimal performance in massive MU-MIMO systems. However, L-MMSE equalization requires accurate knowledge of the signal and noise powers [41]. Furthermore, corresponding hardware designs must solve linear systems of equations, which requires high arithmetic precision and suffers from stringent data dependencies—both of these aspects result in relatively high circuit complexity [32,34,35,139,184]. In addition, practical massive MU-MIMO BS

designs will most likely rely on inexpensive RF circuitry which suffers from numerous impairments, including amplifier nonlinearities, phase noise, and quantization artifacts [40, 185]. The presence of non-ideal hardware necessitates the design of new equalization algorithms that are resilient to real-world hardware imperfections.

Recently, a novel algorithm called *NonParametric Equalizer* (NOPE, for short) was proposed in [41]. NOPE does not require knowledge of the signal and noise powers while provably achieving the performance of L-MMSE equalization in massive MU-MIMO systems. NOPE combines approximate message passing (AMP) [81] with Stein's unbiased risk estimator (SURE) [186] and mismatched data detection [153], which renders this algorithm resilient to numerous hardware impairments while being computationally efficient: NOPE only requires matrix-vector products and avoids a computation of costly matrix inverses or matrix decompositions, which are typically required by L-MMSE equalizer algorithms. Despite all these advantages, NOPE has been designed only for idealistic channel models and has not yet been integrated in hardware.

## 8.2 A Primer on L-MMSE Equalization

We now introduce the system model and review the basics of L-MMSE equalization. We then discuss NOPE.

### 8.2.1 System Model

We consider the input-output relation $\mathbf{y} = \mathbf{H}\mathbf{s}_0 + \mathbf{n}$ to model a massive MU-MIMO uplink system operating in a frequency-flat channel [22]. The vector $\mathbf{y} \in \mathbb{C}^B$ contains the received signals at the BS; $B$ denotes the number of BS antennas; the matrix $\mathbf{H} \in \mathbb{C}^{B \times U}$ represents the uplink MIMO channel; $U$ denotes the number of UEs; the transmit signal vector is $\mathbf{s}_0 \in \mathbb{C}^U$; and the vector $\mathbf{n} \in \mathbb{C}^B$ models receive noise, which has i.i.d. circularly-symmetric complex Gaussian entries with variance $N_0$ per entry. Throughout this chapter of the thesis, we assume that the transmit signal vector $\mathbf{s}_0$ has i.i.d. entries so that $p(\mathbf{s}_0) = \prod_{u=1}^{U} p(s_{0,u})$, where $p(\cdot)$ models the signal prior (e.g., a 16-QAM constellation) with zero mean and signal variance $E_s = \mathbb{E}_{s_{0,u}}\left[|s_{0,u}|^2\right]$, $u = 1, \ldots, U$. We assume that each entry of the channel matrix $\mathbf{H}$ is distributed i.i.d. circularly-symmetric complex Gaussian with variance $1/B$ per complex entry.

### 8.2.2 Basics of L-MMSE Equalization

L-MMSE equalization is among the most popular methods to compute an estimate for $\hat{\mathbf{s}}$ from $\mathbf{y}$ and from knowledge of the channel matrix $\mathbf{H}$, and enjoys widespread use for data detection in MIMO systems [6, 139, 187, 188]. The relatively low computational complexity (except for the inversion of a potentially large matrix) and acceptable performance render this method a feasible alternative to more complicated data detection algorithms. Moreover, it has been shown in [6, 22] that L-MMSE equalization enables (often significantly) higher achievable rates than zero-forcing (ZF) or matched filter (MF)-based equalizers in massive MU-MIMO systems. However, to enable near-optimal spectral efficiency

via L-MMSE equalization, accurate knowledge of the signal and noise powers is required; see, e.g., [13, 41].

Mathematically, the goal of L-MMSE equalization is to compute a linear estimate $\hat{\mathbf{s}} = \mathbf{W}\mathbf{y}$ from the receive vector $\mathbf{y}$ that minimizes the $MSE = E_s\mathbf{s}_0,\mathbf{n}\|\hat{\mathbf{s}} - \mathbf{s}_0\|^2$ using knowledge of the channel matrix $\mathbf{H}$ as well as the signal and noise powers, $E_s$ and $N_0$, respectively. For a circularly-symmetric complex-valued transmit signal $\mathbf{s}_0$, the equalization matrix $\mathbf{W} \in \mathbb{C}^{U\times B}$ is given by $\mathbf{W} = (\mathbf{H}^{\mathrm{H}}\mathbf{H} + \frac{N_0}{E_s}\mathbf{I}_U)^{-1}\mathbf{H}^{\mathrm{H}}$. If the signal $\mathbf{s}_0$ is zero-mean and real-valued (e.g., for BPSK signals), then the optimal linear estimator for the real part $\hat{\mathbf{s}}_{\mathrm{Re}}$ is given by

$$\hat{\mathbf{s}}_{\mathrm{Re}} = \left(\mathbf{H}_{\mathrm{Re}}^{\mathrm{T}}\mathbf{H}_{\mathrm{Re}} + \mathbf{H}_{\mathrm{Im}}^{\mathrm{T}}\mathbf{H}_{\mathrm{Im}} + \frac{N_0}{2E_s}\mathbf{I}_U\right)^{-1}(\mathbf{H}_{\mathrm{Re}}^{\mathrm{T}}\mathbf{y}_{\mathrm{Re}} + \mathbf{H}_{\mathrm{Im}}^{\mathrm{T}}\mathbf{y}_{\mathrm{Im}}),$$

where $\mathbf{H}_{\mathrm{Re}}$, $\mathbf{H}_{\mathrm{Im}}$, $\mathbf{y}_{\mathrm{Re}}$, and $\mathbf{y}_{\mathrm{Im}}$ are the real and imaginary parts of $\mathbf{H}$ and $\mathbf{y}$, respectively; the imaginary part of the estimate is $\hat{\mathbf{s}}_{\mathrm{Im}} = \mathbf{0}_{U\times 1}$. Clearly, L-MMSE equalization relies on knowledge of the quantities $\rho = N_0/E_s$ or $\rho = N_0/(2E_s)$, which requires (i) means to detect whether the transmit signals are real- or complex-valued and (ii) an accurate estimate of $\rho$ that is commonly acquired in a dedicated training phase [189].

### 8.2.3 L-MMSE Equalization via mcB-AMP

As shown in Section 4.3, L-MMSE equalization can be implemented using the mismatched complex Bayesian AMP (mcB-AMP) framework, which we will refer to as M-LAMA-MMSE. By assuming a mismatched Gaussian signal prior distribution $\tilde{p}(\mathbf{s}) = \prod_{u=1}^{U}\tilde{p}(s_u)$ with $\tilde{p}(s_u) \sim \mathcal{CN}(0, E_s)$ instead of the true signal prior $s_0 \sim p(s_0)$ (e.g., two Dirac delta functions concentrated at $-1$ and $+1$ for

BPSK), one can design the following *parametric* L-MMSE algorithm given by M-LAMA in Algorithm 3:

$$\tilde{\sigma}_t^2 = \frac{1}{B}\left\|\mathbf{r}^t\right\|_2^2$$

$$\hat{\tau}_t^2 = \arg\min_{\tau^2 \geq 0} \Psi(\tilde{\sigma}_t^2, \tau^2) \tag{8.1}$$

$$\hat{\mathbf{z}}^t = \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t$$

$$\hat{\mathbf{s}}^{t+1} = \frac{E_s}{E_s + \hat{\tau}_t^2}\hat{\mathbf{z}}^t \tag{8.2}$$

$$\mathbf{r}^{t+1} = \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \beta\mathbf{r}^t\frac{E_s}{E_s + \hat{\tau}_t^2}. \tag{8.3}$$

Interestingly, the estimate $\hat{\mathbf{z}} = \lim_{t\to\infty}\hat{\mathbf{z}}^t$ computed by M-LAMA-MMSE exhibits the same MSE as that of the L-MMSE equalizer in the large-system limit and for matrices $\mathbf{H}$ with uniform channel gains [153]. While this is an asymptotic equivalence, reference [41] has shown that the error-rate performance of M-LAMA-MMSE is virtually indistinguishable from an L-MMSE equalizer in practical (finite-dimensional) massive MU-MIMO systems for a small number of iterations $t_{\max}$ (ten or fewer). Clearly, M-LAMA-MMSE mainly relies on matrix-vector multiplications, which enables parallel hardware designs. However, the exact knowledge of $E_s$ is still required.

## 8.3 Nonparametric Equalizer (NOPE)

We now summarize the necessary steps to free M-LAMA-MMSE from knowledge of the signal power, leading to NOPE. We then propose a generalization of the algorithm that makes it suitable for more realistic MIMO system scenarios.

### 8.3.1 The NOPE Algorithm

To develop NOPE, we wish to automatically tune the signal power $E_s$ and the parameter $\tau_t^2$. To this end, we first introduce the parameter $\kappa^t = E_s/\tau_t^2$ and reparametrize the functions $\mathsf{F}^{\text{mm}}(\hat{z}_u, \kappa^t) = \frac{\kappa^t}{\kappa^t+1}\hat{z}_u$ and $\mathsf{F}'^{\text{mm}}(\hat{z}_u, \kappa^t) = \frac{\kappa^t}{\kappa^t+1}$ in M-LAMA-MMSE. Now, only a single parameter must be tuned per iteration, i.e., $\kappa^t$. Interestingly, [153, Thm. 3] shows that optimal parameter tuning is achieved by tuning each parameter $\kappa^t$ by minimizing (8.1) separately at iteration $t$ starting from $t = 1$ to $t_{\text{max}}$. Hence, the remaining piece is to replace the MSE function $\Psi$ with a function that does not depend on the true signal prior $p(s_0)$. As shown in [41], one can use Stein's unbiased risk estimate (SURE) [186] to extract an estimate of the MSE function $\Psi$ as

$$\hat{\Psi}^{\text{mm}}(\tilde{\sigma}_t^2, \kappa^t) = \tilde{\sigma}_t^2 \frac{\kappa^t - 1}{\kappa^t + 1} + \frac{\|\hat{\mathbf{z}}^t\|_2^2}{U(\kappa^t+1)^2}. \tag{8.4}$$

Since the minimum of $\hat{\Psi}$ is given by $\kappa_{\text{min}}^t = \|\hat{\mathbf{z}}^t\|_2^2/(U\tilde{\sigma}_t^2) - 1$ we can replace the tuning stage in (8.1) by $\kappa_{\text{min}}^t$, which leads to NOPE. As proven in [41, Cor. 6], NOPE achieves the performance of an L-MMSE equalizer in the large antenna limit given that $\mathbf{H}$ has uniform channel gains and for $t \to \infty$.

### 8.3.2 Robust Version of NOPE

NOPE and M-LAMA-MMSE require the matrix $\mathbf{H}$ to have uniform channel gains. However, in practice each UE typically has a different large-scale fading gain (e.g., affected by the distance to the BS), resulting in channel matrices $\mathbf{H}$ whose columns have different scale. We now show how NOPE can be made robust to such channels. As in [41], one can rewrite the channel matrix as $\mathbf{H} = \widetilde{\mathbf{H}}\mathbf{D}$, where each element of $\widetilde{\mathbf{H}}$ is distributed as $\mathcal{CN}(0, 1/B)$ and $\mathbf{D}$ is diagonal containing the

$u$th UE's individual large-scale fading gain $d_u$. For this model, one must estimate the gain of the $u$th UE by $\sum_{b=1}^{B}|H_{b,u}|^2 = d_u^2 \sum_{b=1}^{B}|\tilde{H}_{b,u}|^2$, which converges to $d_u^2$ in the large-antenna limit. Thus, $\mathbf{D}$ is estimated with a diagonal matrix $\widehat{\mathbf{D}}$, where the $u$th diagonal element is given by $\hat{d}_u$. To enable NOPE to support nonuniform channel gains, we modify the posterior mean function in (8.2) into an element-wise operation [41]

$$\mathsf{F}_u^{\mathrm{mm}}(\hat{z}_u^t, \hat{\tau}_t^2) = \frac{E_s}{E_s + \hat{\tau}_t^2/\hat{d}_u^2}\hat{z}_u^t. \tag{8.5}$$

Furthermore, step (8.3) in M-LAMA-MMSE must be replaced by

$$\mathbf{r}^{t+1} = \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \beta\mathbf{r}^t \frac{1}{U}\sum_{u=1}^{U} \mathsf{F}'^{\mathrm{mm}}_u(\hat{z}_u^t, \hat{\tau}_t^2)$$

to take into account the fact that different functions $\mathsf{F}_u^{\mathrm{mm}}(\hat{z}_u^t, \hat{\tau}_t^2)$ are used for each UE. This generalization also requires new estimates for the parameters $E_s$ and $\hat{\tau}_t^2$ in (8.5). As shown in [41, Thm 7], both of these parameters can be estimated as follows

$$\hat{E}_{\mathrm{s}}^t = \frac{v_{\hat{\mathbf{z}}^t} - 2v_{\mathbf{r}}^t}{\sum_{u=1}^{U}\hat{d}_u^2} \quad \text{and} \quad \hat{\tau}_t^2 = \frac{1}{B}\|\mathbf{r}^t\|_2^2, \tag{8.6}$$

where we introduced shorthand notation $v_{\hat{\mathbf{z}}^t} = \sum_{u=1}^{U}\hat{d}_u^2|\hat{z}_u^t|^2$ for the weighted-norm of $\hat{\mathbf{z}}^t$ with respect to its large-scale fading gains, and $v_{\mathbf{r}}^t = \beta\|\mathbf{r}\|_2^2/2$ for the residual norm.

The remaining piece of our robust NOPE is to enable L-MMSE data detection for BPSK constellations for which the imaginary part of $\hat{\mathbf{s}}$ is zero. In fact, assuming a circularly-symmetric complex Gaussian prior for BPSK signals is a poor match as the imaginary part is zero. We generalize NOPE by estimating the signal variance $E_s$ in (8.6) for the real and imaginary parts separately, which enables us to automatically adapt NOPE to the used constellation set. To do so, we

decompose the weighted-norm of $\hat{\mathbf{z}}^t$ denoted by $v_{\hat{\mathbf{z}}^t}$, into real and imaginary parts, i.e., $v_{\hat{\mathbf{z}}^t} = v^t_{\hat{\mathbf{z}},\mathrm{Re}} + v^t_{\hat{\mathbf{z}},\mathrm{Im}}$. More specifically, we can estimate the necessary variances as

$$\hat{E}^t_{\mathrm{s,Re}} = \frac{v^t_{\hat{\mathbf{z}},\mathrm{Re}} - v^t_{\mathbf{r}}}{\sum_{u=1}^{U} \hat{d}_u^2} \quad \text{and} \quad \hat{E}^t_{\mathrm{s,Im}} = \frac{v^t_{\hat{\mathbf{z}},\mathrm{Im}} - v^t_{\mathbf{r}}}{\sum_{u=1}^{U} \hat{d}_u^2},$$

for which $\hat{E}^t_{\mathrm{s,Re}} + \hat{E}^t_{\mathrm{s,Im}} = \hat{E}^t_{\mathrm{s}}$. With all these ingredients, we arrive at the generalized NOPE algorithm in Algorithm 6.

**Algorithm 6** (Robust version of NOPE)**.** *The robust version of the NOPE algorithm proposed in [41] is given by the following iterations: Initialize the algorithm by* $t = 1$, $\hat{\mathbf{s}}^1 = \mathbf{0}_{U\times 1}$, $\mathbf{r}^t = \mathbf{0}$, *and* $\langle \boldsymbol{\alpha}^t \rangle = 0$. *Then, for each iteration* $t = 1, 2, \ldots, t_{\max}$, *compute the following set of equations:*

$$\begin{aligned}
\mathbf{r}^t &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^t + \frac{\beta}{2}\langle \boldsymbol{\alpha} \rangle \mathbf{r}^t \\
v^t_{\mathbf{r}} &= \frac{\beta}{2}\left\|\mathbf{r}^t\right\|_2^2 && (8.7)\\
\hat{\mathbf{z}}^t &= \hat{\mathbf{s}}^t + \hat{\mathbf{d}}^{-2} \circ (\mathbf{H}^{\mathrm{H}}\mathbf{r}^t) && (8.8)\\
v^t_{\hat{\mathbf{z}},\mathrm{Re}} &= \sum_{u=1}^{U} \hat{d}_u^2 \mathrm{Re}\{\hat{z}^t_u\}^2 \\
v^t_{\hat{\mathbf{z}},\mathrm{Im}} &= \sum_{u=1}^{U} \hat{d}_u^2 \mathrm{Im}\{\hat{z}^t_u\}^2 \\
K^t &= (v^t_{\mathbf{r}} \langle \hat{\mathbf{d}}^2 \rangle)^{-1} \\
\alpha^t_{u,\mathrm{Re}} &= (1 + (K^t \hat{d}_u^2 (v^t_{\hat{\mathbf{z}},\mathrm{Re}} - v^t_{\mathbf{r}}))^{-1})^{-1}, \quad u = 1, \ldots, U && (8.9)\\
\alpha^t_{u,\mathrm{Im}} &= (1 + (K^t \hat{d}_u^2 (v^t_{\hat{\mathbf{z}},\mathrm{Im}} - v^t_{\mathbf{r}}))^{-1})^{-1}, \quad u = 1, \ldots, U && (8.10)\\
\hat{\mathbf{s}}^{t+1} &= \boldsymbol{\alpha}^t_{\mathrm{Re}} \circ \mathrm{Re}\big(\hat{\mathbf{z}}^t\big) + \mathrm{i}\boldsymbol{\alpha}^t_{\mathrm{Im}} \circ \mathrm{Im}\big(\hat{\mathbf{z}}^t\big) && (8.11)\\
\boldsymbol{\alpha}^t &= \boldsymbol{\alpha}^t_{\mathrm{Re}} + \boldsymbol{\alpha}^t_{\mathrm{Im}} && (8.12)\\
\rho^t_u &= (2B/\beta) K^t \langle \hat{\mathbf{d}}^2 \rangle \hat{d}_u^2, \quad u = 1, \ldots, U && (8.13)
\end{aligned}$$

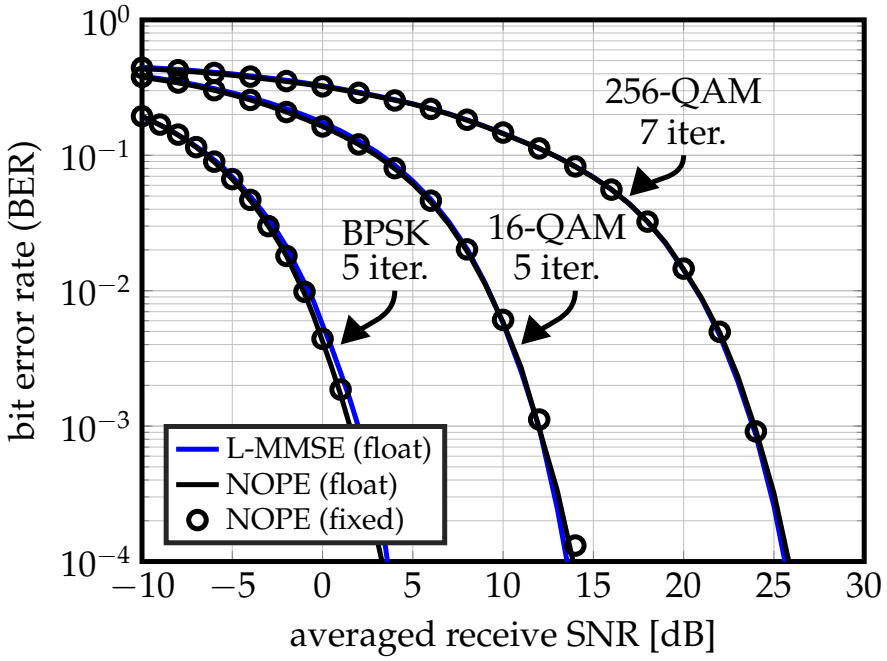



Figure 8.1: Uncoded bit error rate of NOPE algorithm in a $64 \times 16$ massive MU-MIMO system with Rayleigh fading channel matrices. NOPE closely approaches the performance of the L-MMSE estimator that requires exact knowledge of the signal and noise powers. Furthermore, fixed-point arithmetic in NOPE (shown with circle markers) does not exhibit a significant BER performance loss compared to infinite-precision arithmetic (continuous lines).

### 8.3.3 Numerical Results

Figure 8.1 shows uncoded bit error rate (BER) simulation results in a $B = 64$ BS antenna, $U = 16$ UE massive MU-MIMO system with BPSK, 16-QAM, and 256-QAM. We show the performance of exact L-MMSE equalization, as well as the performance of NOPE for both infinite and fixed-precision. Solid lines correspond to floating-point precision, and circle markers correspond to fixed-point precision simulations of NOPE. Evidently, the BER performance of NOPE with $t_{\max} = 5$ iterations ($t_{\max} = 7$ for 256-QAM) is virtually indistinguishable from the exact L-MMSE estimator, which requires accurate knowledge of both the signal and noise powers. Due to its parameter free nature, NOPE is suitable

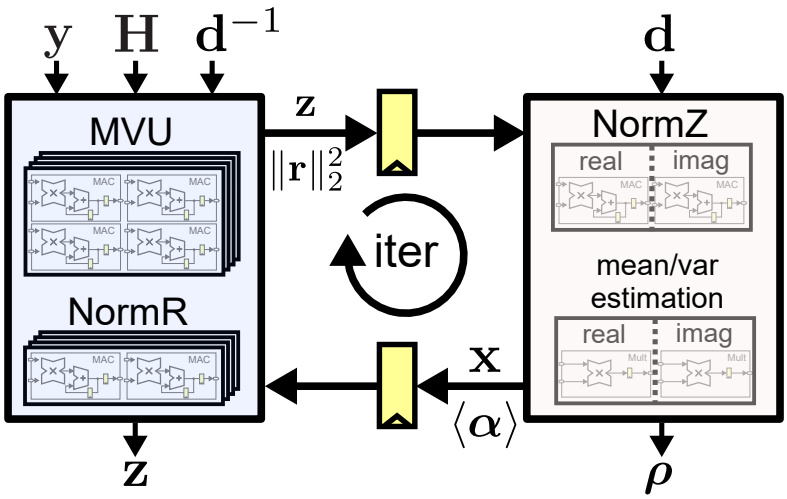



Figure 8.2: Top level block diagram of NOPE algorithm: Each iteration of the NOPE is partitioned into two main units: matrix-vector unit (left) and parameter estimation unit (right). Both units require an identical number of clock cycles, which allows us to process two independent problems via coarse-grained pipeline interleaving.

for situations in which the signal and noise powers change rapidly (e.g., due to interference) or if the transmit constellation is unknown and must be estimated prior to data detection.

## 8.4 VLSI Architecture and Synthesis Results

We now propose a very-large scale integration (VLSI) architecture of the NOPE algorithm for a $B = 64$ BS antenna, $U = 16$ UE massive MU-MIMO system. We then discuss the most essential optimization steps and finally present implementation results in a 28 nm CMOS technology.

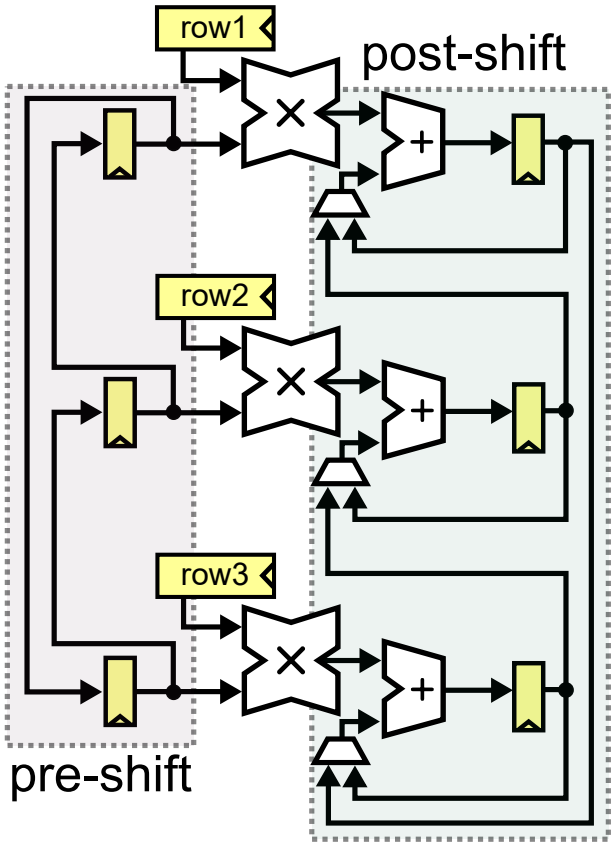



Figure 8.3: Illustration of the matrix-vector unit (MVU) that computes computes $\mathbf{H}\hat{\mathbf{s}}$ and $\mathbf{H}^{\mathrm{H}}\mathbf{r}$ in a $3 \times 3$ system via Cannon's algorithm. To compute $\mathbf{H}\hat{\mathbf{s}}$, we circularly shift the inputs (pre-shift); to compute $\mathbf{H}^{\mathrm{H}}\mathbf{r}$, we circularly shift the outputs (post-shift). This approach enables column-wise storage of the entries of $\mathbf{H}$ without causing access contentions, leading to high throughput.

### 8.4.1 Architecture Overview

We partition the NOPE iterations into two phases, each executed by a separate unit; see Fig. 8.2 for an architecture overview. The *matrix-vector unit* (MVU) executes the necessary matrix-vector multiplications and the *estimation unit* (EU) implements automatic parameter tuning.

The MVU performs the matrix-vector multiplication required to compute the 16 dimensional output vector $\hat{\mathbf{z}}$ in (8.8) and the scalar $\|\mathbf{r}\|_2^2$ in (8.7). The EU implements the mean and variance estimation to compute the posterior mean $\hat{\mathbf{s}}$

in (8.11) and Onsager constant $\langle \boldsymbol{\alpha} \rangle$ in (8.12). In addition, we compute the per-user post-equalization SNR $\boldsymbol{\rho}$ in (8.13), which is required for log-likelihood ratio (LLR) value calculations to perform soft-output data detection. Both units carry out their tasks in the same number of clock cycles, which enables us to process two independent equalization problems concurrently in the same architecture by means of coarse-grained pipeline interleaving.

### 8.4.2 Architecture Details

We now provide architecture details for the MVU and EU, and briefly discuss the key fixed-point implementation aspects.

#### MVU Details

The MVU computes both $\mathbf{H}\hat{\mathbf{s}}$ and $\mathbf{H}^{\mathrm{H}}\mathbf{r}$ in a single unified architecture, similarly to the architecture in [184]. We divide the $64 \times 16$-dimensional channel matrix $\mathbf{H}$ into four $16 \times 16$ blocks, each of which are processed using a separate MVU, which we refer to as MVU-$m$, $m \in \{1, 2, 3, 4\}$. Each $16 \times 16$ matrix-vector multiplication is carried out using 16 complex-valued multiply-accumulate (MAC) units; the matrix-vector operation is carried out on a column-by-column basis so that each MAC unit is associated to a row of the matrix.

A straightforward approach to compute $\mathbf{H}\hat{\mathbf{s}}$ would be to broadcast the 16-dimensional vector $\hat{\mathbf{s}}$ to all MVUs. To compute $\mathbf{H}^{\mathrm{H}}\mathbf{r}$ within the same architecture, access contentions would arise as one would need to be able to read all entries from the row of $\mathbf{H}^{\mathrm{H}}$ and sum all partial products. To enable highly parallel matrix-

vector computation without causing access contentions, we use an architecture as depicted in Fig. 8.3 that performs a variant of Cannon's algorithm [42]. Let $\mathbf{A}$ be a $16 \times 16$ block of $\mathbf{H}$ where each row $r$ is cyclically shifted by its index. To compute $\mathbf{A}\hat{\mathbf{s}}$, the input $\hat{\mathbf{s}}$ is first loaded into the input shift registers (the pre-shift block). We then circularly shift the entries of this shift register while sequentially calculating the MAC operations with entries of the matrix $\mathbf{A}$; the outputs are accumulated in the registers at the output of each MAC unit. This effectively implements a column-by-column matrix-vector operation in 16 clock cycles. To compute $\mathbf{H}^{\mathrm{H}}\mathbf{r}$, we load $\mathbf{r}$ into the input shift register but no cyclical shifts are carried out. Instead we cyclically exchange the outputs (the post-shift block) while accumulating the results. This effectively implements a row-by-row matrix-vector operation in 16 clock cycles.

After the computation of $\mathbf{H}^{\mathrm{H}}\mathbf{r}$, we have to accumulate the results of the four $16 \times 16$ blocks. We do this over two additional clock cycles: in cycle 1, MVU-1 and MVU-4 pass their result to MVU-2 and MVU-3 for accumulation; in cycle 2, MVU-2 passes its result to MVU-3 to obtain the final result.

**EU Details**

The EU computes the posterior mean $\hat{\mathbf{s}}$ and Onsager constant $\langle \boldsymbol{\alpha} \rangle$. To this end, the EU first computes the 16-dimensional norm of the real and imaginary part of $\hat{\mathbf{z}}$, $v_{\hat{\mathbf{z}},\mathrm{Re}}$ and $v_{\hat{\mathbf{z}},\mathrm{Im}}$. We employ two MAC units which compute the real and imaginary parts over 16 clock cycles. Once $v_{\hat{\mathbf{z}},\mathrm{Re}}$ and $v_{\hat{\mathbf{z}},\mathrm{Im}}$ are completed, we compute the so-called denoising parameter $\alpha_{u,\mathrm{Re}}$ in (8.9) and $\alpha_{u,\mathrm{Im}}$ in (8.10) for each $u$th UE sequentially over 16 clock cycles. We note that the function $(1 + x^{-1})^{-1} = 1 - (1 + x)^{-1} \in [0, 1)$ for (8.9) and (8.10) is numerically stable so we

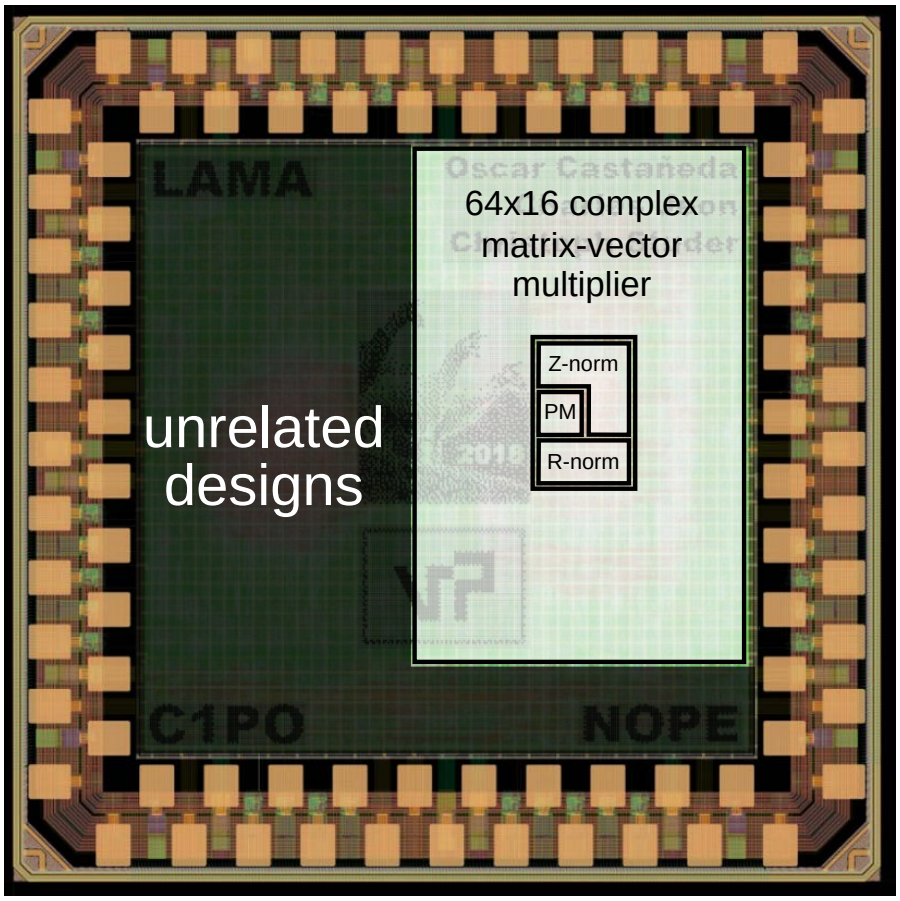



Figure 8.4: Post-placement layout of the NOPE chip with highlighted processing units

employ a single-iteration of the LUT-based Newton-Raphson procedure.

**Fixed-Point Arithmetic**

In order to achieve low hardware complexity and high throughput, our design uses fixed-point arithmetic. We first globally scale $\mathbf{H}$ so that the real and imaginary entries are less than 1. We then quantize each element of $\mathbf{H}$ to 10 fraction bits, and $\mathbf{y}$ to 6 integer and 4 fraction bits. The fixed-point performance of our NOPE design is shown in Fig. 8.1. The solid lines correspond to floating-point performance, the markers to the fixed-point performance of our golden model.

Table 8.1: Synthesis results of NOPE for a 64 BS antenna, 16 UE system and comparison to existing massive MU-MIMO data detectors.

| | This work | Prabhu [34] | Tang [35] | Peng [32] | Castañeda [173] |
|---|---|---|---|---|---|
| System ($B \times U$) | $64 \times 16$ | $128 \times 8$ | $128 \times 32$ | $128 \times 8$ | $128 \times 8$ |
| Algorithm | NOPE | MMSE/ZF[a] | MPD | MMSE | SDR |
| Parameters | none | $E_s$,$N_0$[c] | $E_s$,$N_0$ | $E_s$,$N_0$ | $E_s$,$N_0$ |
| Modulation | 256-QAM | 256-QAM | 256-QAM | 64-QAM | QPSK |
| Preproc. included | no | yes | no | yes | no |
| Preproc. quantities | col. gains | – | Gram mat. | – | Gram mat. |
| Results | synthesis | silicon | silicon | silicon | post-layout |
| Technology [nm] | 28 | 28 | 40 | 65 | 45 |
| Area [mm$^2$] | 0.28 | 1.10 | 0.58 | 2.57 | 0.48 |
| Frequency [MHz] | 800 | 300 | 425 | 680 | 560 |
| Throughput [Gb/s] | 0.92 | 0.30 | 2.76 | 1.02 | 0.13 |
| Eff.[b] [Gb/s/mm$^2$] | 3.29 | 0.27 | 13.87 | 4.96 | 1.08 |

[a]this design also supports precoding; [b]standard technology scaling rules apply; [c]the ZF mode does not require any parameters.

### 8.4.3 Implementation Results and Conclusions

Table 8.1 shows synthesis results for NOPE in a 28 nm CMOS technology and compares our design to existing data detectors for massive MU-MIMO. We note that the numbers reported in Table 8.1 for NOPE are based on synthesis results; an ASIC design is part of ongoing work (see Fig. 8.4 for the post-placement layout of NOPE chip with the highlighted processing units) While our design is comparable to other designs in terms of hardware efficiency, we emphasize that NOPE is completely parameter-free (other than knowledge of $\mathbf{H}$ and $\mathbf{y}$), which makes it more resilient to parameter mismatch and dynamic variations of the system compared to all the other methods. In addition, NOPE requires a minimal amount of preprocessing, i.e., $\hat{\mathbf{d}}^2$ and $\hat{\mathbf{d}}^{-2}$, in contrast to, e.g., the design of [35] that requires computation of the Gram which often dominates the complexity of massive MU-MIMO data detectors [22]. In summary, NOPE is a robust "fire-and-forget" equalization algorithm for MU-MIMO systems that achieves L-MMSE performance at competitive implementation complexity.

# CHAPTER 9

# CONCLUSIONS AND OUTLOOK

Massive multi-user (MU) multiple-input multiple-output (MIMO) will be a core technology in next-generation wireless systems. By equipping the infrastructure base-stations (BSs) with hundreds of antenna elements and serving tens of user equipments (UEs) in the same time-frequency resource, massive MU-MIMO enables orders-of-magnitude higher spectral efficiency than existing wireless systems. The presence of large number of antenna elements at the BS, however, causes significant implementation challenges. In particular, optimal data detection at the BS that maximizes the spectral efficiency (i.e., the number of bits that can be transmitted reliably over a given bandwidth) entails prohibitive complexity. As a result, the majority of existing data detection algorithms for massive MU-MIMO and corresponding hardware designs are sub-optimal, thereby sacrificing spectral efficiency.

In this thesis, we have provided a potential solution for optimal data detection in practical massive MU-MIMO systems. We have tackled the optimal data detection problem by a holistic approach that spans theory, algorithm development, and ASIC design. Concretely, we presented a range of solutions on theory, algorithm, and hardware level that enable near-optimal data detection in practice. In addition, we proposed new methods that reduce the complexity of channel-matrix preprocessing as well as novel architectures and algorithms that enable parallel processing of the most critical tasks in massive MU-MIMO BSs. In order to demonstrate the effectiveness of all our solutions in practically-relevant communication scenarios, we demonstrated our findings via theoretical results, numerical simulations, and ASIC implementations.

There are numerous avenues for future work. We outline the avenues based on three categories (i) theory, (ii) algorithm development, and (iii) VLSI design.

(i) **Theory:** The optimality proof of LAMA was based on the state evolution framework, which was proven rigorously in [82]. Although the asymptotic performance of AMP-based algorithms can be obtained exactly through state evolution, characterizing the performance of the individually optimal data detection is highly non-trivial. References [10, 11, 104] utilized the Replica method [190] from statistical physics to characterize the performance of the individually optimal data detector, which requires additional Replica assumptions. References [12, 78, 113] assumed the exchange on the limits of sparsity and the UEs [113], which we also assumed to prove the optimality of LAMA. The full characterization of optimality of AMP and connections to the results from Replica method, however, is a part of ongoing work [115, 191, 192]. In addition, the performance of AMP for finite dimensions have been presented in [193]. Extending such results for LAMA will provide a framework for a more accurate performance characterization in practical finite-dimensional massive MU-MIMO systems.

(ii) **Algorithm development:** Practical wireless channels do not exhibit i.i.d. Rayleigh fading and has finite number of antennas at both ends of the wireless link. As a result, the optimality of LAMA for practical wireless systems is no longer guaranteed. To mitigate performance loss of LAMA in practice, we proposed damping techniques [124] that provides a good trade-off in performance and complexity. Recent work on expectation-propagation [174] and vector AMP [90] has shown to be robust in practical massive MU-MIMO scenarios, but does so at the cost of additional computational complexity, which is due to the computation of inverse of the

Gram matrix. Therefore, the development of novel methods that not only are robust in practical wireless channels, but also are of low computational complexity is left for future work.

(iii) **VLSI design:** The majority of the detector implementations for massive MU-MIMO systems are for small number of UEs [32, 34, 173]. Although there exist two 32-UE detectors [35, 36], they employ a sub-optimal detection algorithm and either only support hard outputs or support QPSK constellation only. In addition, the detector exhibits error flooring behavior in practical wireless channels. There has been a recent implementation of a 16-UE detector [37] that employs expectation-propagation [174] which is robust in practical, correlated channels. Since the LAMA implementation in this thesis is the first near-optimal soft-input soft-output 32-UE detector, additional 32-UE detector implementations for a more accurate comparison are left for future work.

APPENDIX A

**PROOFS**

## A.1 Proofs for Chapter 3

### A.1.1 Proof of [1, Lem. 5.56] for complex-valued systems

For completeness, we include [1, Lem. 5.56] and its proof for complex-valued MIMO systems. We will use Lemma 44 to derive the cB-AMP algorithm in Appendix A.1.2.

**Lemma 44.** *Let $\hat{s}^t_{u\to b}$ and $\hat{\tau}^2_{u\to b,t}$ be the mean and variance of the distribution $\nu^t_{u\to b}$ in (A.3), respectively. Suppose at iteration $t$, the messages are set to $\hat{v}^t_{b\to u}(s_u) = \hat{\phi}^t_{b\to u}(s_u)$, where $\hat{\phi}^t_{b\to u}(s_u)$ is defined by*

$$\hat{\phi}^t_{b\to u}(s_u) \triangleq \frac{|H_{b,u}|^2}{\pi(N_0^{\text{post}} + \hat{\tau}^2_{b\to u,t})} \exp\left(-\frac{\left|H_{b,u}s_u - r^t_{b\to u}\right|^2}{(N_0^{\text{post}} + \hat{\tau}^2_{b\to u,t})}\right), \tag{A.1}$$

*and $\hat{\tau}^2_{b\to u,t} = \hat{\tau}^2_t$, where the residual and variance terms are given by*

$$r^t_{b\to u} \triangleq y_b - \sum_{\tilde{u}\neq u} H_{b,\tilde{u}}\hat{s}_{\tilde{u}\to b}, \quad \hat{\tau}^2_{b\to u,t} \triangleq \sum_{\tilde{u}\neq u} |H_{b,\tilde{u}}|^2 \hat{\tau}^2_{\tilde{u}\to b,t}.$$

*Then, at the next iteration $t+1$, the mean and the variance of the message $\nu^{t+1}_{u\to b}$ are given by*

$$\hat{s}^{t+1}_{u\to b} = \mathsf{F}\left(\sum_{\tilde{b}\neq b} H^*_{\tilde{b},u} r^t_{\tilde{b}\to u}, N_0^{\text{post}} + \hat{\tau}^2_t\right),$$

$$\hat{\tau}^2_{u\to b,t+1} = \mathsf{G}\left(\sum_{\tilde{b}\neq b} H^*_{a,u} r^t_{\tilde{b}\to u}, N_0^{\text{post}} + \hat{\tau}^2_t\right),$$

*Proof.* Suppose at iteration $t$, the messages from factor nodes to the variable nodes are set to be $\hat{v}^t_{b\to u} = \hat{\phi}^t_{b\to u}$. Then,

$$\begin{aligned}
\nu^{t+1}_{u\to b} &= \prod_{\tilde{b}\neq b} \hat{\phi}^t_{\tilde{b}\to u} p(s_u) \\
&= \exp\left( - \frac{\sum_{\tilde{b}\neq b} |H_{\tilde{b},u} s_u - r^t_{\tilde{b}\to u}|^2}{N_0^{\text{post}} + \hat{\tau}_t^2} \right) p(s_u) \\
&= \frac{1}{Z} \exp\left( - \frac{|s_u|^2 - 2\sum_{\tilde{b}\neq b} \mathrm{Re}\left( s_u^* H^*_{\tilde{b},u} r^t_{\tilde{b}\to u} \right)}{N_0^{\text{post}} + \hat{\tau}_t^2} \right) \exp\left( \frac{|s_u|^2}{B(N_0^{\text{post}} + \hat{\tau}_t^2)} \right) p(s_u) \\
&= \phi^{t+1}_{u\to b}(s_u) \left\{ 1 + O(|s_u|^2 / B) \right\},
\end{aligned}$$

where $Z$ is a normalization constant that ensures $\nu^{t+1}_{u\to b}$ is a probability density function. Here, we defined $\phi^{t+1}_{u\to k}$ as

$$\phi^{t+1}_{u\to b}(s_u) = f\left( s_u \Big| \sum_{\tilde{b}\neq b} H^*_{\tilde{b},u} r^t_{\tilde{b}\to u} \right), \tag{A.2}$$

where $f(s_u|\hat{s}_u)$ is the posterior distribution of an Gaussian distribution defined in (3.4). Here, the noise variance of the Gaussian distribution corresponds to $N_0^{\text{post}} + \hat{\tau}_t^2$. By definition, the mean $\mathsf{F}$ and variance $\mathsf{G}$ of $\nu^{t+1}_{u\to k}$ is given as mean and variance of the conditional probability distribution defined in (A.2):

$$\begin{aligned}
\hat{s}^{t+1}_{u\to b} &= \mathsf{F}\left( \sum_{\tilde{b}\neq b} H^*_{\tilde{b},u} r^t_{\tilde{b}\to u}, N_0^{\text{post}} + \hat{\tau}_t^2 \right), \\
\hat{\tau}^2_{u\to b,t+1} &= \mathsf{G}\left( \sum_{\tilde{b}\neq b} H^*_{\tilde{b},u} r^t_{\tilde{b}\to u}, N_0^{\text{post}} + \hat{\tau}_t^2 \right).
\end{aligned}$$

□

### A.1.2 Derivation of Algorithm 1

We start by considering a factor graph $G = (\mathcal{U}, \mathcal{B}, E)$ with variable nodes $\mathcal{U} = \{1, \ldots, U\}$, factor nodes $\mathcal{B} = \{1, \ldots, B\}$, and edges $E = \mathcal{U} \times \mathcal{B} =$

$\{(u,b) : u \in \mathcal{U}, b \in \mathcal{B}\}$. The sum-product message equations for (3.3) at every iteration $t$ are given by [105],

$$\nu^t_{u\to b}(s_u) = \prod_{\tilde{b}\neq b} \hat{\nu}^{t-1}_{\tilde{b}\to u}(s_u)\, p(s_u), \tag{A.3}$$

$$\hat{\nu}^t_{b\to u}(s_u) = \int_{\mathbb{C}} p(y_b \,|\, \mathbf{s}) \prod_{\tilde{u}\neq u} \nu^t_{\tilde{u}\to b}(s_{\tilde{u}})\, \mathrm{d}y_b, \tag{A.4}$$

where $\nu^t_{u\to b}(s_u)$ and $\hat{\nu}^t_{b\to u}(s_u)$ are probability density functions.

Now, with Lemma 44 we can simplify the sum-product algorithm shown in (A.3) and (A.4). We first expand the messages $\hat{s}^{t+1}_{u\to b}$ and $r^{t+1}_{b\to u}$ into two parts (i) constant messages $\hat{s}^{t+1}_u$ and $r^{t+1}_b$ which are independent of the edge $(u,b)$ and (ii) perturbed messages $\Delta\hat{s}^{t+1}_{u\to b}, \Delta r^{t+1}_{b\to u}$ that depend on the edge. As done in [9, Eq. 5], we assume $\Delta\hat{s}^{t+1}_{u\to b}, \Delta r^{t+1}_{b\to u} = O(1/\sqrt{U})$ such that

$$\hat{s}^{t+1}_{u\to b} \triangleq \hat{s}^{t+1}_u + \Delta\hat{s}^{t+1}_{u\to b} + O(1/U), \tag{A.5}$$

$$r^{t+1}_{b\to u} \triangleq r^{t+1}_b + \Delta r^{t+1}_{b\to u} + O(1/U). \tag{A.6}$$

We then replace the complex-valued soft-thresholding function $\eta(\cdot)$ by the conditional mean $\mathsf{F}(\cdot)$ as in [9, Prop. II.1], and use the decomposition in (A.5) and (A.6) to obtain the following:

$$r^t_{b\to u} = y_b - \sum_{\tilde{u}\neq u} H_{b,\tilde{u}}\hat{s}_{\tilde{u}\to b}, = y_b - \sum_{\tilde{u}=1}^{U} H_{b,\tilde{u}}(\hat{s}^{t+1}_{\tilde{u}} + \Delta\hat{s}^{t+1}_{\tilde{u}\to b}) + H_{b,u}\hat{s}^t_u + O(1/U).$$

We will let $r^{t+1}_b = y_b - \sum_{u=1}^{U} H_{b,u}(\hat{s}^{t+1}_u + \Delta\hat{s}^{t+1}_{u\to b})$ and $\Delta r^{t+1}_{b\to u} = H_{b,u}\hat{s}^t_u$. For $\hat{s}^{t+1}_{u\to b}$,

we obtain:

$$
\begin{aligned}
\hat{s}_{u\to b}^{t+1} &= \mathsf{F}\Bigg( \sum_{\tilde{b}\neq b} H_{\tilde{b},u}^{*} r_{\tilde{b}\to u}^{t}, N_0^{\text{post}} + \hat{\tau}_t^2 \Bigg) \\
&= \mathsf{F}\Bigg( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*} (r_{\tilde{b}}^{t} + \Delta r_{\tilde{b}\to u}^{t}) - H_{b,u}^{*} r_b^t, N_0^{\text{post}} + \hat{\tau}_t^2 \Bigg) + O(1/U) \\
&\overset{(a)}{=} \mathsf{F}\Bigg( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*} (r_{\tilde{b}}^{t} + \Delta r_{\tilde{b}\to u}^{t}), N_0^{\text{post}} + \hat{\tau}_t^2 \Bigg) \\
&\quad - \mathrm{Re}\big(H_{b,u}^{*} r_b^t\big)\, \partial_1 \mathsf{F}^{\mathrm{R}} \Bigg( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*} (r_{\tilde{b}}^{t} + \Delta r_{\tilde{b}\to u}^{t}) \Bigg) \\
&\quad - \mathrm{Im}\big(H_{b,u}^{*} r_b^t\big)\, \partial_2 \mathsf{F}^{\mathrm{R}} \Bigg( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*} (r_{\tilde{b}}^{t} + \Delta r_{\tilde{b}\to u}^{t}) \Bigg) \\
&\quad - \mathrm{Re}\big(H_{b,u}^{*} r_b^t\big)\, \partial_1 \mathsf{F}^{\mathrm{I}} \Bigg( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*} (r_{\tilde{b}}^{t} + \Delta r_{\tilde{b}\to u}^{t}) \Bigg) \\
&\quad - \mathrm{Im}\big(H_{b,u}^{*} r_b^t\big)\, \partial_2 \mathsf{F}^{\mathrm{I}} \Bigg( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*} (r_{\tilde{b}}^{t} + \Delta r_{\tilde{b}\to u}^{t}) \Bigg) \\
&\quad + O(1/U),
\end{aligned}
$$

where for $(a)$, we follow the first-order Taylor expansion [9, Eq. (15)] with shorthand notations:

$$
\begin{aligned}
\partial_1 \mathsf{F}^{\mathrm{R}} &\triangleq \frac{\partial \mathrm{Re}\big(\mathsf{F}(x+\mathrm{i}y, \tau^2)\big)}{\partial x}, \quad \partial_2 \mathsf{F}^{\mathrm{R}} \triangleq \frac{\partial \mathrm{Re}\big(\mathsf{F}(x+\mathrm{i}y, \tau^2)\big)}{\partial y}, \\
\partial_1 \mathsf{F}^{\mathrm{I}} &\triangleq \frac{\partial \mathrm{Im}\big(\mathsf{F}(x+\mathrm{i}y, \tau^2)\big)}{\partial x}, \quad \partial_2 \mathsf{F}^{\mathrm{I}} \triangleq \frac{\partial \mathrm{Im}\big(\mathsf{F}(x+\mathrm{i}y, \tau^2)\big)}{\partial y}.
\end{aligned}
$$

Similar to $r_b^{t+1}$, we let $\hat{s}_u^{t+1} = \mathsf{F}\left( \sum_{b=1}^{B} H_{b,u}^{*}(r_b^t + \Delta r_{\tilde{b}\to u}^t), N_0^{\text{post}} + \hat{\tau}_t^2 \right)$ and

$$\begin{aligned}
\Delta\hat{s}_{u\to b}^{t+1} = &-\text{Re}\big(H_{b,u}^{*} r_b^t\big)\, \partial_1 \mathsf{F}^{\text{R}}\left( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*}(r_{\tilde{b}}^t + \Delta r_{\tilde{b}\to u}^t) \right) \\
&- \text{Im}\big(H_{b,u}^{*} r_b^t\big)\, \partial_2 \mathsf{F}^{\text{R}}\left( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*}(r_{\tilde{b}}^t + \Delta r_{\tilde{b}\to u}^t) \right) \\
&- \text{Re}\big(H_{b,u}^{*} r_b^t\big)\, \partial_1 \mathsf{F}^{\text{I}}\left( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*}(r_{\tilde{b}}^t + \Delta r_{\tilde{b}\to u}^t) \right) \\
&- \text{Im}\big(H_{b,u}^{*} r_b^t\big)\, \partial_2 \mathsf{F}^{\text{I}}\left( \sum_{\tilde{b}=1}^{B} H_{\tilde{b},u}^{*}(r_{\tilde{b}}^t + \Delta r_{\tilde{b}\to u}^t) \right).
\end{aligned}$$

Now that we have defined these quantities, we perform to simplify them. First of all, we observe that:

$$\begin{aligned}
\hat{s}_u^{t+1} &= \mathsf{F}\left( \sum_{b=1}^{B} H_{b,u}^{*}(r_{\tilde{b}}^t + \Delta r_{b\to u}^t), N_0^{\text{post}} + \hat{\tau}_t^2 \right) \\
&= \mathsf{F}\left( \sum_{b=1}^{B} H_{b,u}^{*} r_b^t + \sum_{b=1}^{B} |H_{b,u}|^2 \hat{s}_u^t, N_0^{\text{post}} + \hat{\tau}_t^2 \right) \\
&= \mathsf{F}\left( \hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^{*} r_b^t, N_0^{\text{post}} + \hat{\tau}_t^2 \right). \qquad \text{(A.7)}
\end{aligned}$$

Similarly, the message variance term in the large-system limit converges to:

$$\begin{aligned}
\hat{\tau}_{u\to b,t+1}^2 &= \mathsf{G}\left( \sum_{\tilde{b}\neq b} H_{\tilde{b},u}^{*} r_{\tilde{b}\to u}^t, N_0^{\text{post}} + \hat{\tau}_t^2 \right) \\
&= \mathsf{G}\left( \hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^{*} r_b^t, N_0^{\text{post}} + \hat{\tau}_t^2 \right).
\end{aligned}$$

We note that the message-variance update equation was not provided in [9] and hence, we include it for completeness. Second, we observe the residual term $r_b^{t+1}$ can be rewritten as:

$$\begin{aligned}
r_b^{t+1} &= y_b - \sum_{u=1}^{U} H_{b,u}(\hat{s}_u^{t+1} + \Delta\hat{s}_{u\to b}^{t+1}) \\
&= y_b - \sum_{u=1}^{U} H_{b,u}\hat{s}_u^{t+1} - \sum_{u=1}^{U} H_{b,u}\Delta\hat{s}_{u\to b}^{t+1}.
\end{aligned}$$

Now that we have all the terms ready, we proceed to obtaining cB-AMP. We start by simplifying the residual update by plugging in $\Delta\hat{s}_{u\to b}^{t+1}$.

$$\begin{aligned} r_b^{t+1} = y_b &- \sum_{u=1}^{U} H_{b,u}\hat{s}_u^{t+1} \\ &+ \sum_{u=1}^{U} H_{b,u}\mathrm{Re}\big(H_{b,u}^* r_b^t\big)\,\partial_1\mathsf{F}^{\mathrm{R}}\left(\hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^* r_b^t\right) \\ &+ \sum_{u=1}^{U} H_{b,u}\mathrm{Im}\big(H_{b,u}^* r_b^t\big)\,\partial_2\mathsf{F}^{\mathrm{R}}\left(\hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^* r_b^t\right) \\ &+ \sum_{u=1}^{U} H_{b,u}\mathrm{Re}\big(H_{b,u}^* r_b^t\big)\,\partial_1\mathsf{F}^{\mathrm{I}}\left(\hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^* r_b^t\right) \\ &+ \sum_{u=1}^{U} H_{b,u}\mathrm{Im}\big(H_{b,u}^* r_b^t\big)\,\partial_2\mathsf{F}^{\mathrm{I}}\left(\hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^* r_b^t\right) \end{aligned} \tag{A.8}$$

We simplify the residual update by using the fact that $\mathbf{H}$ satisfies (A1). Since the columns of $\mathbf{H}$ have unit norm with pairwise independence by (A1), each term in (A.8) can be simplified in the large system limit as follows:

$$\begin{aligned} \sum_{u=1}^{U} H_{b,u}\partial_1\mathsf{F}^{\mathrm{R}}\mathrm{Re}\big(H_{b,u}^* r_b^t\big) &= \frac{\beta}{2} r_b^t\langle\partial_1\mathsf{F}^{\mathrm{R}}\rangle, \\ \sum_{u=1}^{U} H_{b,u}\partial_2\mathsf{F}^{\mathrm{R}}\mathrm{Im}\big(H_{b,u}^* r_b^t\big) &= \frac{\beta}{2\mathrm{i}} r_b^t\langle\partial_2\mathsf{F}^{\mathrm{R}}\rangle, \\ \mathrm{i}\sum_{u=1}^{U} H_{b,u}\partial_1\mathsf{F}^{\mathrm{I}}\mathrm{Re}\big(H_{b,u}^* r_b^t\big) &= \frac{\beta\mathrm{i}}{2} r_b^t\langle\partial_1\mathsf{F}^{\mathrm{I}}\rangle, \\ \mathrm{i}\sum_{u=1}^{U} H_{b,u}\partial_2\mathsf{F}^{\mathrm{R}}\mathrm{Im}\big(H_{b,u}^* r_b^t\big) &= \frac{\beta}{2} r_b^t\langle\partial_2\mathsf{F}^{\mathrm{I}}\rangle, \end{aligned}$$

Finally, the next iteration message variance update is computed in the large-system limit by:

$$\begin{aligned} \hat{\tau}_{t+1}^2 = \sum_{u=1}^{U} |H_{b,u}|^2\,\hat{\tau}_{u\to b,t+1}^2 &= \sum_{u=1}^{U} \frac{1}{B}\mathsf{G}\left(\hat{s}_u^t + \sum_{b=1}^{B} H_{b,u}^* r_b^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2\right). \\ &= \beta\langle\mathsf{G}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2)\rangle \end{aligned} \tag{A.9}$$

By using the Hadamard product, we arrive at Algorithm 1.

### A.1.3 Proof of Lemma 1

We use the facts that $\partial_2 \mathsf{F}^{\mathrm{R}}$, $\partial_1 \mathsf{F}^{\mathrm{I}}$, and $\partial_2 \mathsf{F}^{\mathrm{I}}$ are all zero for real-valued systems. Moreover, since the $u_2$-norm of each column of $\mathbf{H}$ is one according to (A1), the update (A.8) simplifies to

$$\begin{aligned} r_b^{t+1} - \left(y_b - \mathbf{h}_b^{\mathrm{row}} \hat{\mathbf{s}}^{t+1}\right) &= r_b^t \sum_{u=1}^{U} H_{b,u}^2 \mathsf{F}' \left( \hat{s}_u^t + \sum_{b=1}^{B} H_{b,u} r_b^t, \hat{\tau}_t^2 \right) \\ &\simeq \beta r_b^t \left\langle \mathsf{F}' \left( \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{T}} \mathbf{r}^t, \hat{\tau}_t^2 \right) \right\rangle, \end{aligned} \tag{A.10}$$

where $\mathsf{F}'$ is the derivative of the mean function $\mathsf{F}(\hat{s}_u, \hat{\tau}^2)$ taken with respect to $\hat{s}_u$. The final comparison of (A.10) with [1, Eq. 5.74] reveals equivalence of real-valued cB-AMP and B-AMP.

### A.1.4 Intuitive derivation of Theorem 2

We present a non-rigorous derivation of Theorem 2 for complex-valued systems; a rigorous proof can be found in [82]. Assume that the MIMO channel $\mathbf{H}(t)$ changes each iteration $t$, where the elements are distributed $\mathcal{CN}(0, 1/B)$. In addition, let $\mathsf{F}(z, \tau^2)$ and $\mathsf{G}(z, \tau^2)$ are functions defined in (3.5) and (3.6) according to the mean and variance of the distribution in (3.4), respectively. Let $\mathbf{y}^t = \mathbf{H}(t)\mathbf{s}_0 + \mathbf{n}$ where the entries of $\mathbf{n}$ are circularly symmetric complex Gaussian with variance $N_0$. Assuming that we fix the postulated noise variance to $N_0^{\mathrm{post}}$, then, in each iteration, the recursion is defined as:

$$\mathbf{r}^t = \mathbf{y}^t - \mathbf{H}\hat{\mathbf{s}}^t, \tag{A.11}$$

$$\hat{\mathbf{s}}^{t+1} = \mathsf{F}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}(t)\mathbf{r}^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2), \tag{A.12}$$

$$\hat{\tau}_{t+1}^2 = \beta \langle \mathsf{G}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}(t)\mathbf{r}^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2) \rangle. \tag{A.13}$$

By substituting $\mathbf{r}^t$ in (A.11) into $\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}(t)\mathbf{r}^t$, we have that

$$\begin{aligned}\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}(t)\mathbf{r}^t &= \mathbf{H}^{\mathrm{H}}(t)\mathbf{y}^t + \Big(\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}(t)\mathbf{H}(t)\Big)\hat{\mathbf{s}}^t \\ &= \mathbf{s}_0 + \mathbf{H}^{\mathrm{H}}(t)\mathbf{n} + \Big(\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}(t)\mathbf{H}(t)\Big)\big(\hat{\mathbf{s}}^t - \mathbf{s}_0\big). \end{aligned} \tag{A.14}$$

The central limit theorem shows that each diagonal and non-diagonal entry in $\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}(t)\mathbf{H}(t)$ is distributed $\mathcal{N}(0, 1/B)$ and $\mathcal{CN}(0, 1/B)$ respectively, with pairwise independent entries. Also, for each $u$th entry in $\big(\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}(t)\mathbf{H}(t)\big)\ (\hat{\mathbf{s}}^t - \mathbf{s}_0)$, the real and imaginary parts are normally distributed with zero mean and variance $\frac{\|\hat{\mathbf{s}}^t - \mathbf{s}_0\|_2^2}{2B} + \delta_u^t$ and $\frac{\|\hat{\mathbf{s}}^t - \mathbf{s}_0\|_2^2}{2B} - \delta_u^t$ respectively, with $\delta_u^t = \frac{\mathrm{Re}\{(\hat{s}_u^t - s_{0,u})^2\}}{2B}$. With

$$\hat{\sigma}_t^2 = \lim_{U\to\infty} \big\|\hat{\mathbf{s}}^t - \mathbf{s}_0\big\|^2 / U, \tag{A.15}$$

and noting that $\delta_u^t \to 0$ as $U \to \infty$, we have that

$$\Big(\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}(t)\mathbf{H}(t)\Big)\big(\hat{\mathbf{s}}^t - \mathbf{s}_0\big) \to \mathcal{CN}\Big(0, \beta\hat{\sigma}_t^2\Big).$$

Moreover, by conditioning on $\mathbf{n}$, $\mathbf{H}^{\mathrm{H}}(t)\mathbf{n} \to \mathcal{CN}(0, N_0)$ by the law of large numbers. By Definition 3 of the effective noise variance of cB-AMP, we have the relation $\sigma_t^2 = N_0 + \beta\hat{\sigma}_t^2$ with $\hat{\sigma}_t^2$ defined in (A.15). Thus, each $u$th entry of (A.12) converges to $\mathsf{F}(s_{0u} + \sigma_t N_{\mathbb{C}}, N_0^{\mathrm{post}} + \hat{\tau}_t^2)$ where $N_{\mathbb{C}} \sim \mathcal{CN}(0, 1)$. Since we assume a fixed prior distribution for all $s_{0,u}$, we obtain the following recursion for (A.12):

$$\begin{aligned}\sigma_{t+1}^2 &= N_0 + \beta \lim_{U\to\infty} \frac{1}{U}\big\|\hat{\mathbf{s}}^{t+1} - \mathbf{s}_0\big\|^2 \\ &= N_0 + \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\Big[\big|\mathsf{F}(S + \sigma_t N_{\mathbb{C}}, N_0^{\mathrm{post}} + \tau_t^2) - S\big|^2\Big],\end{aligned}$$

with $S \sim p(S)$ and we introduced $\tau_t^2 = \lim_{U\to\infty} \hat{\tau}_t^2$. Starting from (A.13), we use (A.14) and (A.15), and the law of large numbers to obtain:

$$\begin{aligned}\tau_{t+1}^2 &= \beta \lim_{U\to\infty} \langle \mathsf{G}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}(t)\mathbf{r}^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2)\rangle \\ &= \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\big[\mathsf{G}(S + \sigma_t N_{\mathbb{C}}, N_0^{\mathrm{post}} + \tau_t^2)\big].\end{aligned}$$

By introducing the postulated variance $\gamma_t^2 = N_0^{\text{post}} + \tau_t^2$, we obtain the final cSE:

$$\sigma_{t+1}^2 = N_0 + \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\left[\left|\mathsf{F}(S + \sigma_t N_{\mathbb{C}}, \gamma_t^2) - S\right|^2\right]$$
$$\gamma_{t+1}^2 = N_0^{\text{post}} + \beta\, \mathbb{E}_{S,N_{\mathbb{C}}}\Big[\mathsf{G}(S + \sigma_t N_{\mathbb{C}}, \gamma_t^2)\Big].$$

We reiterate that the formulation of $\mathbf{H}(t)$ to obtain cSE in Theorem 2 was non-rigorous; a rigorous proof can be found in [82].

### A.1.5 Proof of Lemma 4

We start with cB-AMP as detailed in Algorithm 1. We simplify intermediate steps in cB-AMP using the definition of $\mathsf{F}(\hat{s}_u, \tau^2)$ and $\mathsf{G}(\hat{s}_u, \tau^2)$, and our knowledge of the prior distribution. Recall that $\mathsf{F}(\hat{s}_u, \tau^2)$ in (3.19) was defined as

$$\mathsf{F}(\hat{s}_u, \tau^2) = \sum_{a \in \mathcal{O}} w_a(\hat{s}_u, \tau^2) a,$$

By taking partial derivatives of $\mathsf{F}(\hat{s}_u, \tau^2)$ with the notations defined in Algorithm 1, we have the following expressions, where we drop the notation $w_a = w_a(\hat{s}_u, \tau^2)$ for simplicity.

$$\partial_1 \mathsf{F}^{\mathrm{R}} = \frac{2}{\tau^2}\left[\sum_{a \in \mathcal{O}} \mathrm{Re}(a)^2\, w_a - \left(\sum_{a \in \mathcal{O}} \mathrm{Re}(a)\, w_a\right)^2\right],$$
$$\partial_2 \mathsf{F}^{\mathrm{I}} = \frac{2}{\tau^2}\left[\sum_{a \in \mathcal{O}} \mathrm{Im}(a)^2\, w_a - \left(\sum_{a \in \mathcal{O}} \mathrm{Im}(a)\, w_a\right)^2\right],$$
$$\partial_2 \mathsf{F}^{\mathrm{I}} = \partial_1 \mathsf{F}^{\mathrm{R}}$$
$$= \frac{2}{\tau^2}\left[\sum_{a \in \mathcal{O}} \mathrm{Re}(a)\,\mathrm{Im}(a)\, w_a - \Big(\sum_{a \in \mathcal{O}} \mathrm{Re}(a)\, w_a\Big)\Big(\sum_{a \in \mathcal{O}} \mathrm{Im}(a)\, w_a\Big)\right].$$

Note that (3.20) can be separated in real and imaginary parts. Therefore,

$$\begin{aligned}
\mathsf{G}(\hat{s}_u, \tau^2) &= \sum_{a\in\mathcal{O}} |a|^2 \, w_a - \left| \sum_{a\in\mathcal{O}} a w_a \right|^2 \\
&= \sum_{a\in\mathcal{O}} \mathrm{Re}(a)^2 \, w_a - \left( \sum_{a\in\mathcal{O}} \mathrm{Re}(a) \, w_a \right)^2 \\
&\quad + \sum_{a\in\mathcal{O}} \mathrm{Im}(a)^2 \, w_a - \left( \sum_{a\in\mathcal{O}} \mathrm{Im}(a) \, w_a \right)^2 \\
&= \frac{\tau^2}{2} \Big[ \partial_1 \mathsf{F}^{\mathrm{R}} + \partial_2 \mathsf{F}^{\mathrm{I}} \Big] (\hat{s}_u, \tau^2)
\end{aligned}$$

Finally, observe that $\partial_1 \mathsf{F}^{\mathrm{I}} = \partial_2 \mathsf{F}^{\mathrm{R}}$ and $1/i = -i$, so Algorithm 1 simplifies to:

$$\begin{aligned}
\hat{\mathbf{s}}^{t+1} &= \mathsf{F}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}} \mathbf{r}^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2) \\
\hat{\tau}_{t+1}^2 &= \beta \langle \mathsf{G}(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}} \mathbf{r}^t, N_0^{\mathrm{post}} + \hat{\tau}_t^2) \rangle \\
\mathbf{r}^{t+1} &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \frac{\hat{\tau}_{t+1}^2}{N_0^{\mathrm{post}} + \hat{\tau}_t^2} \mathbf{r}^t.
\end{aligned}$$

### A.1.6 Proof of Lemma 5

Since $N_0^{\mathrm{post}} \to \infty$, the recursions in Algorithm 2 are given by

$$\begin{aligned}
\hat{\mathbf{s}}^t &= \lim_{N_0^{\mathrm{post}} \to \infty} \mathsf{F}\Big( \hat{\mathbf{s}}^{t-1} + \mathbf{H}^{\mathrm{H}} \mathbf{r}^{t-1}, N_0^{\mathrm{post}} + \hat{\tau}_t^2 \Big), \\
\mathbf{r}^t &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^t.
\end{aligned}$$

First of all, notice that as $N_0^{\mathrm{post}} \to \infty$, $w_a(\hat{s}_u, N_0^{\mathrm{post}} + \hat{\tau}_t^2) \to p_a$ for any $\hat{s}_u$. Therefore, for all $t$,

$$\begin{aligned}
\hat{\mathbf{s}}^t &= \lim_{N_0^{\mathrm{post}} \to \infty} \mathsf{F}\Big( \hat{\mathbf{s}}^{t+1} + \mathbf{H}^{\mathrm{H}} \mathbf{r}^{t-1}, N_0^{\mathrm{post}} + \hat{\tau}_t^2 \Big) \to \sum_{a\in\mathcal{O}} a p_a = 0, \\
\mathbf{r}^t &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^t = \mathbf{y},
\end{aligned}$$

and thus, the Gaussian output $\hat{\mathbf{z}}^t = \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t$ is equivalent to the matched filter output $\mathbf{H}^{\mathrm{H}}\mathbf{y}$. To show that the non-linear MMSE output corresponds to the matched filter involves computing $\lim_{N_0^{\text{post}}\to\infty} \frac{N_0^{\text{post}}}{E_s}\hat{\mathbf{s}}^{t+1}$, which is given by

$$\begin{aligned}
\lim_{N_0^{\text{post}}\to\infty} \frac{N_0^{\text{post}}}{E_s}\hat{\mathbf{s}}^{t+1} &= \lim_{N_0^{\text{post}}\to\infty} \frac{N_0^{\text{post}}}{E_s}\mathsf{F}\left(\mathbf{H}^{\mathrm{H}}\mathbf{y}, N_0^{\text{post}} + \hat{\tau}_t^2\right) \\
&\overset{(a)}{=} \lim_{N_0^{\text{post}}\to\infty} \frac{N_0^{\text{post}}}{E_s} \sum_{a\in\mathcal{O}} a p_a \left(1 - \frac{1}{N_0^{\text{post}} + \hat{\tau}_t^2}\left|\mathbf{H}^{\mathrm{H}}\mathbf{y} - a\right|^2\right) \\
&= -\frac{1}{E_s}\sum_{a\in\mathcal{O}} a p_a \left|\mathbf{H}^{\mathrm{H}}\mathbf{y} - a\right|^2 = \mathbf{H}^{\mathrm{H}}\mathbf{y},
\end{aligned}$$

where $(a)$ follows from expansion of $\exp(\cdot)$ function in $w_a(\hat{s}_u, N_0^{\text{post}} + \hat{\tau}_t^2)$, and the absolute value is taken element-wise.

### A.1.7 Proof of Lemma 6

First, note that as $\beta \to 0$ for a fixed $N_0^{\text{post}}$, we have that $\hat{\tau}_t^2 = 0$ for all $t \geq 1$. Therefore, we have the following recursions,

$$\begin{aligned}
\hat{\mathbf{s}}^{t+1} &= \mathsf{F}\left(\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t, N_0^{\text{post}}\right), \\
\mathbf{r}^{t+1} &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1}.
\end{aligned}$$

Following the derivation of complex state evolution in Appendix A.1.4, as $\beta \to 0$, we have

$$\hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t = \mathbf{H}^{\mathrm{H}}\mathbf{y} + (\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}\mathbf{H})\hat{\mathbf{s}}^t \to \mathbf{H}^{\mathrm{H}}\mathbf{y},$$

because the entries of $(\mathbf{I}_U - \mathbf{H}^{\mathrm{H}}\mathbf{H})\hat{\mathbf{s}}^t$ converge to a complex normal distribution with zero mean and variance $\beta\tilde{\sigma}_t^2$ with $\tilde{\sigma}_t^2 = \lim_{U\to\infty} \frac{1}{U}\left\|\hat{\mathbf{s}}^t\right\|^2$. Since $\tilde{\sigma}_t^2$ is finite and $\beta \to 0$, we have that the Gaussian output of LAMA is $\hat{\mathbf{z}}^t = \mathbf{H}^{\mathrm{H}}\mathbf{y}$ (independent of the iteration index $t$). Hence, the non-linear MMSE output $\hat{\mathbf{s}}^{t+1}$ of LAMA is given by $\mathsf{F}\left(\mathbf{H}^{\mathrm{H}}\mathbf{y}, N_0^{\text{post}}\right)$ for all $t$.

We show that one iteration of LAMA is sufficient to achieve AWGN performance by the cSE in Theorem 2. Recall that previous paragraph demonstrated that $\hat{\mathbf{z}}^t = \mathbf{H}^{\mathrm{H}}\mathbf{y}$ for all $t$. Thus, the equivalent output noise variance is computed as $\sigma^2 = N_0 + \beta\,\mathbb{V}\mathrm{ar}_S[S] = N_0$, where the last step comes from $\beta \to 0$. Since each output of LAMA is identical every iteration and the output noise variance is $N_0$, one iteration is sufficient to achieve AWGN performance.

### A.1.8 Proof of Lemma 9

Note that since $N_0 = 0$, we have $\sigma^2 = \gamma^2$ by Corollary 3. Since the variance of $S$ is finite, denote $\mathbb{V}\mathrm{ar}_S[S] = \sigma_s^2$. By [194, Prop. 15], we have the following upper bound for $\Psi(\sigma^2, \sigma^2)$:

$$\Psi(\sigma^2, \sigma^2) \leq \frac{\sigma_s^2}{\sigma_s^2 + \sigma^2}\sigma^2, \tag{A.16}$$

where equality is achieved for all $\sigma^2$ if and only if $S$ is complex normal with variance $\sigma_s^2$. Note that if $\sigma^2 = 0$, then (A.16) is achieved for any $\sigma_s^2$. If $\sigma^2 > 0$, then

$$\Psi(\sigma^2, \sigma^2) \leq \frac{\sigma_s^2}{\sigma_s^2 + \sigma^2}\sigma^2 = \frac{1}{1 + \sigma^2/\sigma_s^2}\sigma^2 < \sigma^2,$$

and, hence, the proof follows.

The first part of Lemma 9 is trivial from (A.16), and thus, $\Psi(\sigma^2, \sigma^2) \to 0$ as $\sigma^2 \to 0$. The second part is noting that as $\sigma^2 \to \infty$, $\mathsf{F}(\cdot, \sigma^2) \to \sum_{a \in \mathcal{O}} a p_a = \mathbb{E}_S[S]$, and hence we have

$$\lim_{\sigma^2 \to \infty} \Psi(\sigma^2, \sigma^2) \to \mathbb{E}_S\left[|S - \mathbb{E}_S[S]|^2\right] = \mathbb{V}\mathrm{ar}_S[S].$$

### A.1.9 Proof of Theorem 10

We assume the initialization in Algorithm 2. Since $N_0 = N_0^{\text{post}} = 0$, if LAMA perfectly recovers the true signal $\mathbf{s}_0$, then the fixed-point (3.30) is unique at $\sigma^2 = 0$. This happens if the system ratio is strictly less than the ERT, $\beta_{\mathcal{O}}^{\max}$ because otherwise, i.e., $\beta \geq \beta_{\mathcal{O}}^{\max}$, there exists a non-unique fixed point to (3.30) for some $\sigma^2 > 0$ by Definition 6.

### A.1.10 Proof of Lemma 11

We show that for a fixed constellation $\mathcal{O}$, $\beta_{\mathcal{O}}^{\min} \leq \beta_{\mathcal{O}}^{\max}$. For conciseness, define $\sigma_\star^2$ as the fixed-point $\sigma^2 = \beta_{\mathcal{O}}^{\max}\Psi(\sigma^2, \sigma^2)$. The proof is straightforward as,

$$\begin{aligned}\beta_{\mathcal{O}}^{\min} &\overset{(a)}{=} \min_{\sigma^2>0}\left\{\left(\frac{\mathrm{d}\Psi(\sigma^2,\sigma^2)}{\mathrm{d}\sigma^2}\right)^{-1}\right\} \leq \left(\frac{\mathrm{d}\Psi(\sigma^2,\sigma^2)}{\mathrm{d}\sigma^2}\right)^{-1}\Bigg|_{\sigma^2=\sigma_\star^2} \\ &\overset{(b)}{=} \left(\frac{1}{\beta_{\mathcal{O}}^{\max}}\right)^{-1} = \beta_{\mathcal{O}}^{\max},\end{aligned}$$

where (a) and (b) follow from the definitions of MRT and ERT, respectively.

### A.1.11 Proof of Lemma 12

As $\beta < \beta_{\mathcal{O}}^{\max}$, there exists a value of $N_0$, denote it as $N_0^\star$, such that for $N_0 < N_0^\star$, the fixed-point solution of LAMA is unique. We note that $N_0^{\min}(\beta)$ is also a candidate for $N_0^\star$ as the fixed-point solution of LAMA is unique for all $N_0 < N_0^{\min}(\beta)$. In addition, since $\mathcal{O}$ is a constellation, by [108, Thm. 10], $\Psi(\sigma^2, \sigma^2)$ has a continuous derivative and $\lim_{\sigma^2\to 0}\frac{\mathrm{d}}{\mathrm{d}\sigma^2}\Psi(\sigma^2,\sigma^2) = 0$. Hence, there exists a value $\sigma_\star^2$ such that

for all $\sigma^2 < \sigma_\star^2$,

$$\frac{\mathrm{d}}{\mathrm{d}\sigma^2}\Psi(\sigma^2,\sigma^2) < \frac{1}{2\beta}. \tag{A.17}$$

Now, suppose that $N_0^\star < \sigma_\star^2/2$. Then, for all $\sigma^2 < \sigma_\star^2$ we have:

$$N_0 + \beta\Psi(\sigma^2,\sigma^2) \overset{(a)}{<} N_0 + \frac{\sigma^2}{2},$$

where $(a)$ follows from (A.17) and the mean value theorem. Since $2N_0 < 2N_0^\star < \sigma_\star^2$, we have that:

$$N_0 + \beta\Psi(2N_0,2N_0) < N_0 + N_0 = 2N_0,$$

and therefore, the fixed-point solution $\sigma^2$ has to be between $N_0$ and $2N_0$. As a result, as $N_0 \to 0$, the fixed-point solution $\sigma^2 \to 0$. The last part is apparent as:

$$\begin{aligned}\lim_{N_0\to 0} 1 &= \lim_{N_0\to 0}\frac{N_0}{\sigma^2} + \beta\lim_{N_0\to 0}\frac{\Psi(\sigma^2,\sigma^2)}{\sigma^2}\\ &= \lim_{N_0\to 0}\frac{N_0}{\sigma^2} + \beta\lim_{\sigma^2\to 0}\frac{\mathrm{d}}{\mathrm{d}\sigma^2}\Psi(\sigma^2,\sigma^2) = \lim_{N_0\to 0}\frac{N_0}{\sigma^2},\end{aligned}$$

### A.1.12 Proof of Theorem 13

Since the constellation $\mathcal{O}$ is separable, we introduce a shorthand notation for $\mathcal{O}^{\mathrm{R}} = \mathrm{Re}(\mathcal{O})$ and $\mathcal{O}^{\mathrm{I}} = \mathrm{Im}(\mathcal{O})$. It is easy to observe that the weight scalar $w_a(\hat{s}_u,\tau^2)$ can be rewritten as a product between the weight scalar of the real and imaginary constellation $w_{a_{\mathrm{R}}}(\hat{s}_u,\tau^2)$ and $w_{a_{\mathrm{I}}}(\hat{s}_u,\tau^2)$, i.e.,

$$w_a(\hat{s}_u,\tau^2) = w_{a_{\mathrm{R}}}(\hat{s}_u,\tau^2)w_{a_{\mathrm{I}}}(\hat{s}_u,\tau^2),$$

where

$$w_{a_{\mathrm{R}}}(\hat{s}_u,\tau^2) = \frac{p_{a_{\mathrm{R}}}\exp\Big(-\frac{1}{\tau^2}(\mathrm{Re}(\hat{s}_u)-a_{\mathrm{R}})^2\Big)}{\sum_{b\in\mathcal{O}^{\mathrm{R}}} p_b\exp\Big(-\frac{1}{\tau^2}(\mathrm{Re}(\hat{s}_u)-b)^2\Big)}, \tag{A.18}$$

and likewise for $w_{a_{\mathrm{I}}}(\hat{s}_u, \tau^2)$. Therefore, $\mathsf{F}$ is separable because

$$\begin{aligned}
\mathsf{F}(\hat{s}_u, \tau^2) &= \sum_{a \in \mathcal{O}} w_a(\hat{s}_u, \tau^2) a \\
&= \sum_{a \in \mathcal{O}} w_a(\hat{s}_u, \tau^2) a_{\mathrm{R}} + i \sum_{a \in \mathcal{O}} w_a(\hat{s}_u, \tau^2) a_{\mathrm{I}} \\
&= \sum_{a_{\mathrm{R}} \in \mathcal{O}^{\mathrm{R}}} a_{\mathrm{R}} \sum_{a_{\mathrm{I}} \in \mathcal{O}^{\mathrm{I}}} w_a(\hat{s}_u, \tau^2) + i \sum_{a_{\mathrm{I}} \in \mathcal{O}^{\mathrm{I}}} a_{\mathrm{I}} \sum_{a_{\mathrm{R}} \in \mathcal{O}^{\mathrm{R}}} w_a(\hat{s}_u, \tau^2) \\
&= \sum_{a_{\mathrm{R}} \in \mathcal{O}^{\mathrm{R}}} w_{a_{\mathrm{R}}}(\hat{s}_u, \tau^2) a_{\mathrm{R}} + i \sum_{a_{\mathrm{I}} \in \mathcal{O}^{\mathrm{I}}} w_{a_{\mathrm{I}}}(\hat{s}_u, \tau^2) a_{\mathrm{I}}. \qquad \text{(A.19)}
\end{aligned}$$

Now, for a real-valued constellation $\mathcal{O}^{\mathrm{R}}$ for a real-valued system, the message mean $\mathsf{F}^{\mathrm{R}}$ is given by:

$$\mathsf{F}^{\mathrm{R}}(\hat{s}_u, \tau^2) = \sum_{a \in \mathcal{O}^{\mathrm{R}}} w_a^{\mathrm{R}}(\hat{s}_u, \tau^2) a, \qquad \text{(A.20)}$$

where the weight scalar for the real-valued system is computed by

$$w_a^{\mathrm{R}}(\hat{s}_u, \tau^2) = \frac{p_a \exp\left(-\frac{1}{2\tau^2}(\hat{s}_u - a)^2\right)}{\sum_{b \in \mathcal{O}^{\mathrm{R}}} p_b \exp\left(-\frac{1}{2\tau^2}(\hat{s}_u - b)^2\right)} = w_{a_{\mathrm{R}}}(\hat{s}_u, 2\tau^2).$$

Therefore, we have that for $S_{\mathrm{R}} = \mathrm{Re}(S)$, $S_{\mathrm{I}} = \mathrm{Im}(S)$:

$$\begin{aligned}
\Psi(\sigma^2, \gamma^2) &= \mathbb{E}_{S,N}\left[\left|\mathsf{F}(S + \sigma N, \gamma^2) - S\right|^2\right] \\
&\overset{(a)}{=} \mathbb{E}_{S_{\mathrm{R}}, N_{\mathbb{R}}}\left[\left(\mathrm{Re}\left(\mathsf{F}\left(S_{\mathrm{R}} + \frac{\sigma}{\sqrt{2}} N_{\mathbb{R}}, \gamma^2\right)\right) - S_{\mathrm{R}}\right)^2\right] \\
&\quad + \mathbb{E}_{S_{\mathrm{I}}, N_{\mathbb{R}}}\left[\left(\mathrm{Im}\left(\mathsf{F}\left(S_{\mathrm{I}} + \frac{\sigma}{\sqrt{2}} N_{\mathbb{R}}, \gamma^2\right)\right) - S_{\mathrm{I}}\right)^2\right] \\
&\overset{(b)}{=} 2\,\mathbb{E}_{S_{\mathrm{R}}, N_{\mathbb{R}}}\left[\left(\mathrm{Re}\left(\mathsf{F}\left(S_{\mathrm{R}} + \frac{\sigma}{\sqrt{2}} N_{\mathbb{R}}, \gamma^2\right)\right) - S_{\mathrm{R}}\right)^2\right] \\
&\overset{(c)}{=} 2\,\mathbb{E}_{S_{\mathrm{R}}, N_{\mathbb{R}}}\left[\left(\sum_{a \in \mathcal{O}^{\mathrm{R}}} w_a^{\mathrm{R}}\left(S_{\mathrm{R}} + \frac{\sigma}{\sqrt{2}} N_{\mathbb{R}}, \frac{\gamma^2}{2}\right) a - S_{\mathrm{R}}\right)^2\right] \\
&= 2\Psi^{\mathrm{R}}\left(\frac{\sigma^2}{2}, \frac{\sigma^2}{2}\right),
\end{aligned}$$

where $(a)$ follows from (A.19), $(b)$ from definition of separable constellation, and $(c)$ follows from construction of (A.20). We note that the case for variance function $\Phi$ is derived similarly.

### A.1.13 Proof of Lemma 14

We show that for a separable constellation $\mathcal{O}$, the MRT and ERT are equivalent. We start by denoting the complex-valued MSE function as $\Psi(\sigma^2, \sigma^2) = \mathbb{E}_{S,N_{\mathbb{C}}}\left[\left|\mathsf{F}(S + \sigma N_{\mathbb{C}}, \sigma^2) - S\right|^2\right]$, where $N \sim \mathcal{CN}(0,1)$, and $S \sim p(S)$ for constellation $\mathcal{O}$. Denote the real-valued MSE function

$$\Psi^{\mathrm{R}}(\sigma^2, \sigma^2) = \mathbb{E}_{S_{\mathrm{R}},N_{\mathbb{R}}}\left[\left(\mathsf{F}\left(S_{\mathrm{R}} + \sigma N_{\mathbb{R}}, \sigma^2\right) - S_{\mathrm{R}}\right)^2\right],$$

where $N_{\mathbb{R}} \sim \mathcal{N}(0,1)$ and $S_{\mathrm{R}} \sim p(S_{\mathrm{R}})$ for the real-valued constellation $\mathrm{Re}(\mathcal{O})$. We know from Theorem 13 that $\Psi(\sigma^2, \sigma^2) = 2\Psi^{\mathrm{R}}(\sigma^2/2, \sigma^2/2)$. By denoting $\beta_{\mathbb{C}}^{\min}$ and $\beta_{\mathbb{R}}^{\min}$ as the MRT of complex- and real-valued MSE function, respectively, we have:

$$\begin{aligned}\beta_{\mathbb{C}}^{\min} &= \min_{\sigma^2 \geq 0}\left\{\left(\frac{\mathrm{d}\Psi(\sigma^2, \sigma^2)}{\mathrm{d}\sigma^2}\right)^{-1}\right\} = \min_{\sigma^2 \geq 0}\left\{\left(\frac{\mathrm{d}\Psi^{\mathrm{R}}(\frac{\sigma^2}{2}, \frac{\sigma^2}{2})}{\mathrm{d}\sigma^2/2}\right)^{-1}\right\} \\ &= \min_{\bar{\sigma}^2 \geq 0}\left\{\left(\frac{\mathrm{d}\Psi^{\mathrm{R}}(\bar{\sigma}^2, \bar{\sigma}^2)}{\mathrm{d}\bar{\sigma}^2}\right)^{-1}\right\} = \beta_{\mathbb{R}}^{\min}\end{aligned}$$

The remaining quantities, $\beta^{\max}$ and the critical noise levels of Lemma 14 can be derived similarly.

### A.1.14 Proof of Lemma 15

Note that if $\beta < \beta_{\mathcal{O}}^{\min}$, then the slope of the function $\beta\Psi(\sigma^2, \sigma^2)$ with respect to $\sigma^2$ is always less than 1, i.e., $\beta \frac{\mathrm{d}}{\mathrm{d}\sigma^2}\Psi(\sigma^2, \sigma^2)\Big|_{\sigma^2=\sigma_\star^2} < 1$ for any $\sigma_\star^2 > 0$. In addition, since $\beta < \beta_{\mathcal{O}}^{\min}$, the fixed-point solution to $N_0 + \beta\Psi(\sigma^2, \sigma^2) = \sigma^2$ is unique. Now the exponentially-fast convergence result can be shown by using the bounding technique of [1, Lem. 6.4.1].

In [1], the proof for showing exponentially-fast convergence of standard AMP to its largest fixed-point solution was shown by analyzing the stability constant $\mathrm{SC}(\Psi)$ which was defined by:

$$\mathrm{SC}(\Psi) = \beta \frac{\mathrm{d}}{\mathrm{d}\sigma^2}\Psi(\sigma^2, \sigma^2)\Big|_{\sigma^2=\sigma_\star^2},$$
$$\sigma_\star^2 = \max_{\sigma^2>0}\Big\{\sigma^2 : N_0 + \beta\Psi(\sigma^2, \sigma^2) \geq \sigma^2\Big\}$$

Using the new notation of stability constant, the condition of $\mathrm{SC}(\Psi) < 1$, i.e., slope at the largest fixed point is less than 1, was only needed in [1] to show exponential-fast convergence. However, we note that this approach was viable in [1] due to concavity of the MSE function of $\Psi(\sigma^2, \sigma^2)$ for the soft-thresholding function; however, the MSE function of LAMA does not have such properties. In fact, the MSE function for LAMA for commonly used constellation in wireless is neither convex nor concave. However, as shown above, if $\beta < \beta_{\mathcal{O}}^{\min}$, we have that not only $\mathrm{SC}(\Psi) < 1$, but also $\beta \frac{\mathrm{d}}{\mathrm{d}\sigma^2}\Psi(\sigma^2, \sigma^2)\Big|_{\sigma^2=\sigma_\star^2} < 1$ for any $\sigma_\star^2 > 0$. Therefore, if $\beta < \beta_{\mathcal{O}}^{\min}$, LAMA converges exponentially fast to its unique fixed-point solution $\sigma_\star^2$.

## A.2 Proofs for Chapter 4

### A.2.1 Proof of Lemma 17

Since there is no prior mismatch, we note conditional mean under $\sigma_t^2 = \gamma_t^2$ minimizes the MSE, where the MSE is equivalent to the conditional variance [194]. As a result,

$$\min_{\gamma^2} \Psi(\sigma_t^2, \gamma^2) = \mathbb{E}_{S_0, N_{\mathbb{C}}} \left[ \left| \mathsf{F}(S_0 + \sigma_t N_{\mathbb{C}}, \sigma_t^2) - S_0 \right|^2 \right].$$

Therefore, (4.6) and (4.7) are equivalent and reduces to the SE recursion given by cB-AMP in Theorem 2.

### A.2.2 Proof of Lemma 20

The proof follows similarly as the steps in [136] to show that $\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2)$ is quasi-convex in $\gamma_t^2$. To show the quasi-convexity, we will show that $\frac{\mathrm{d}}{\mathrm{d}\gamma_t^2}\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2)$ has only one sign-change. The proof is straightforward as

$$\frac{\mathrm{d}}{\mathrm{d}\gamma_t^2}\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2) = 2\left(\frac{E_s}{E_s + \gamma_t^2}\right)^3 (\gamma_t^2 - \sigma_t^2),$$

so $\frac{\mathrm{d}}{\mathrm{d}\gamma_t^2}\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2)$ has one sign-change at $\gamma_t^2 = \sigma_t^2$. Note that $\frac{\mathrm{d}}{\mathrm{d}\gamma_t^2}\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2)\Big|_{\gamma_t^2 \to 0} < 0$ so $\gamma_t^{\star 2} = \sigma_t^2$ is the global minimizer for $\Psi^{\mathrm{mm}}(\sigma_t^2, \gamma_t^2)$.

### A.2.3 Proof of Lemma 21

We will compute (4.15) by first computing the mismatched SE recursion for $M$-PAM system (under real-valued noise) with equally likely priors. We first

derive the following function $\mathsf{F}^{\alpha}(s_u)$ by

$$\mathsf{F}^{\alpha}(s_u) = \begin{cases} -\alpha & s_u \leq -\alpha, \\ s_u & -\alpha < s_u \leq \alpha, \\ \alpha & s_u > \alpha. \end{cases}$$

Note that under equally likely priors, the $M$-PAM constellation can be expressed by $p(s_u) = \frac{1}{M}\sum_{k=-M/2+1}^{M/2} \delta(s_u - (2k-1))$. Then, for a given $\sigma$ and $N_{\mathbb{R}} \sim \mathcal{N}(0,1)$, for some symbol $2k-1$ we have the following:

$$\begin{aligned} &\mathbb{E}_{N_{\mathbb{R}}}[(\mathsf{F}^{\alpha}((2k-1)+\sigma N_{\mathbb{R}}) - (2k-1))^2] \\ &= \sigma^2 + (\bar{\alpha}_k^2 - \sigma^2)Q\left(\frac{\bar{\alpha}_k}{\sigma}\right) + (\alpha_k^2 - \sigma^2)Q\left(\frac{\alpha_k}{\sigma}\right) \\ &\quad - \frac{\sigma}{\sqrt{2\pi}}\bar{\alpha}_k \exp\left(-\frac{\bar{\alpha}_k^2}{2\sigma^2}\right) - \frac{\sigma}{\sqrt{2\pi}}\alpha_k \exp\left(-\frac{\alpha_k^2}{2\sigma^2}\right), \end{aligned}$$

where we denote $\bar{\alpha}_k = \alpha - (2k-1)$, and $\alpha_k = \alpha + (2k-1)$. Thus, by using symmetry for $M$-PAM symbols:

$$\begin{aligned} \Psi_{\text{PAM}}^{\text{mm}}(\sigma^2) &= \frac{2}{M}\sum_{k=1}^{M/2} \mathbb{E}_{N_{\mathbb{R}}}[(\mathsf{F}^{\alpha}((2k-1)+\sigma N_{\mathbb{R}}) - (2k-1))^2] \\ &= \sigma^2 + \frac{2}{M}\sum_{k=1}^{M/2}\Bigg[(\bar{\alpha}_k^2 - \sigma^2)Q\left(\frac{\bar{\alpha}_k}{\sigma}\right) + (\alpha_k^2 - \sigma^2)Q\left(\frac{\alpha_k}{\sigma}\right) \\ &\quad - \frac{\sigma}{\sqrt{2\pi}}\bar{\alpha}_k \exp\left(-\frac{\bar{\alpha}_k^2}{2\sigma^2}\right) - \frac{\sigma}{\sqrt{2\pi}}\alpha_k \exp\left(-\frac{\alpha_k^2}{2\sigma^2}\right)\Bigg]. \end{aligned} \tag{A.21}$$

To obtain $\Psi_{\text{QAM}}^{\text{mm}}(\sigma^2)$, we use the result from Corollary 18, i.e., $\Psi_{\text{QAM}}^{\text{mm}}(\sigma^2) = 2\Psi_{\text{PAM}}^{\text{mm}}(\sigma^2/2)$.

### A.2.4 Proof of Lemma 23

By (4.15), we can compute $\frac{\mathrm{d}\Psi^{\mathrm{mm}}_{\mathrm{QAM}}(\sigma^2)}{\mathrm{d}\sigma^2}$ directly by:

$$\begin{aligned}\frac{\mathrm{d}\Psi^{\mathrm{mm}}_{\mathrm{QAM}}(\sigma^2)}{\mathrm{d}\sigma^2} = 1 - \frac{1}{M} - \frac{2}{M}\left(\frac{2\alpha}{\sqrt{\pi\sigma^2}}e^{-\frac{4\alpha}{\sigma^2}} + Q\left(\frac{2\alpha}{\sigma/\sqrt{2}}\right)\right) \\ - \frac{2}{M}\sum_{k=1}^{M/2-1}\left[\frac{1}{\sqrt{\pi\sigma^2}}\left(\alpha_k \exp\left(-\frac{\alpha_k^2}{\sigma^2}\right) + \bar{\alpha}_k \exp\left(-\frac{\bar{\alpha}_k^2}{\sigma^2}\right)\right)\right. \\ \left. + Q\left(\frac{\bar{\alpha}_k}{\sigma/\sqrt{2}}\right) + Q\left(\frac{\alpha_k}{\sigma/\sqrt{2}}\right)\right]\end{aligned}$$

Note that all terms right side of $1 - 1/M$ is negative for $\sigma^2 > 0$ and attains maximum of 0 as $\sigma^2 \to 0$. Thus, we have that $\min_{\sigma^2 \geq 0}\left\{\left(\frac{\mathrm{d}\Psi(\sigma^2)}{\mathrm{d}\sigma^2}\right)^{-1}\right\} = \lim_{\sigma^2 \to 0}\left(\frac{\mathrm{d}\Psi(\sigma^2)}{\mathrm{d}\sigma^2}\right)^{-1} = (1 - 1/M)^{-1}$.

To show that M-LAMA also recovers original signal when $\beta = \beta_{\mathcal{O}}^{\max}$, we use the fact that $\frac{\mathrm{d}\Psi(\sigma^2)}{\mathrm{d}\sigma^2}$ is maximized only at $\sigma^2 \to 0$ and hence no other $\sigma_\star^2 > 0$ satisfies $\beta_{\mathcal{O}}^{\min} = (\mathrm{d}\Psi(\sigma_\star^2)/\mathrm{d}\sigma_\star^2)^{-1}$, so the fixed point is unique.

### A.2.5 Proof of Lemma 25

We start with the following result from [14] that establishes the error-rate performance of BOX detector.

**Theorem 45** (Thm 3.1 [14])**.** *Assume a real-valued M-PAM system with* $\beta < (1 - 1/M)^{-1}$*. The symbol-error rate in the large-system limit converges to* $2(1 - 1/M)Q(1/\sigma_\star)$*, where* $\sigma_\star$ *is the unique minimizer to* $F_M(\sigma)$*:*

$$F_M(\sigma) = \frac{\sigma}{2}\left(\frac{1}{\beta} - \frac{M-1}{M}\right) + \frac{N_0}{2\beta\sigma} + \frac{1}{M}\sum_{k \in K} S(\sigma, k), \tag{A.22}$$

*where* $K = \{2, 4, \ldots, 2(M-1)\}$, *and*

$$S(\sigma, k) = \left(\sigma + \frac{k^2}{\sigma}\right) Q\left(\frac{k}{\sigma}\right) - \frac{k}{\sqrt{2\pi}} \exp\left(-\frac{k^2}{2\sigma^2}\right).$$

Compared to the exact expression in [14], we have additional $\beta$ term in the denominator of $\frac{N_0}{2\beta\sigma}$ due to our definition of $\mathsf{SNR} = \beta \frac{E_s}{N_0} = \frac{\beta}{N_0}$. We now show that the minimizer $\sigma_\star$ of (A.22) coincides exactly to that fixed point solution given by state evolution.

Since $\sigma_\star$ is the unique minimal solution to $F_M(\sigma)$, $F'_M(\sigma_\star) = 0$ where $F'_M(\sigma) = \frac{\mathrm{d}}{\mathrm{d}\sigma} F_M(\sigma)$. Straightforward differentiation of $F_M(\sigma)$ gives the following equation:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sigma} F_M(\sigma) = \frac{1}{2}\left(\frac{1}{\beta} - \frac{M-1}{M}\right) - \frac{N_0}{2\beta\sigma^2} \\ + \frac{1}{M} \sum_{k \in K} \left[\left(1 - \frac{k^2}{\sigma^2}\right) Q\left(\frac{k}{\sigma}\right) + \frac{k}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{k^2}{2\sigma^2}\right)\right].\end{aligned}$$

Rearranging $F'_M(\sigma_\star) = 0$ gives

$$\sigma_\star^2 = N_0 + \beta \left[\frac{M-1}{M} \sigma_\star^2 + \frac{2}{M} \sum_{k \in K} T(\sigma_\star, k)\right], \tag{A.23}$$

where we define a shorthand notation for $T(\sigma, k)$:

$$T(\sigma, k) = (k^2 - \sigma^2) Q\left(\frac{k}{\sigma}\right) - \frac{k\sigma}{\sqrt{2\pi}} \exp\left(-\frac{k^2}{2\sigma^2}\right)$$

We now show that (A.23) corresponds to fixed-point solution to the SE equation $\sigma_\star^2 = N_0 + \beta \Psi^{\mathrm{mm}}_{\mathrm{PAM}}(\sigma_\star^2)$, where $\Psi^{\mathrm{mm}}_{\mathrm{PAM}}(\sigma^2)$ is derived in (A.21). We start by partitioning $K = K_{\mathrm{L}} \cup K_M \cup K_{\mathrm{U}}$ where $K_{\mathrm{L}} = \{2, 4, \ldots, M-2\}$, $K_{\mathrm{M}} = M$, and $K_{\mathrm{U}} = \{M+2, \ldots, 2(M-1)\}$. We will use the fact that $T(\sigma, 0) = -\frac{1}{2}\sigma^2$. For $K_{\mathrm{L}}$, we have

$$\begin{aligned}\sum_{k \in K_{\mathrm{L}}} T(\sigma_\star, k) = \sum_{k' \in K_{\mathrm{L}}} T(\sigma_\star, M - k') = \sum_{u=1}^{\frac{M}{2}-1} T(\sigma_\star, M - 2u) \\ = \sum_{u=1}^{M/2} T(\sigma_\star, M - 2u) + \frac{1}{2}\sigma_\star^2.\end{aligned}$$

For $K_{\mathrm{U}}$, we have

$$\sum_{k\in K_{\mathrm{U}}} T(\sigma_\star, k) = \sum_{k'\in K_{\mathrm{L}}} T(\sigma_\star, M+k') = \sum_{u=1}^{\frac{M}{2}-1} T(\sigma_\star, M+2u)$$
$$= \sum_{u=2}^{M/2} T(\sigma_\star, M+2(u-1)),$$

so that

$$\sum_{k\in K_{\mathrm{U}}} T(\sigma_\star, k) + T(\sigma_\star, M) = \sum_{u=1}^{M/2} T(\sigma_\star, M+2(u-1)).$$

Therefore, the proof is complete as the RHS of (A.23) is:

$$\frac{M-1}{M}\sigma_\star^2 + \frac{2}{M}\sum_{k\in K} T(\sigma_\star, k) = \sigma_\star^2$$
$$+ \frac{2}{M}\sum_{k=1}^{M/2}[T(\sigma_\star, \alpha+1-2k) + T(\sigma_\star, \alpha+1+2(k-1))]$$
$$= \sigma_\star^2 + \frac{2}{M}\sum_{k=1}^{M/2}[T(\sigma_\star, \bar{\alpha}_k) + T(\sigma_\star, \alpha_k)] = \Psi_{\mathrm{PAM}}^{\mathrm{mm}}(\sigma_\star^2).$$

We note that the case for BPSK, i.e., $M = 2$, was shown in [15] and the corresponding proof for M-LAMA was given in [153]. The presented proof shows that the BOX-relaxed method in [14] and SM-LAMA under uniform hypercube prior achieves the same fixed-point in (4.10). Moreover, due to the decoupling property of LAMA, the symbol error-rate of real-valued $M$-PAM system is given by $2(1-M^{-1})Q(1/\sigma_\star)$. We note that by Corollary 18, our result can be generalized to that of a $M^2$-QAM systems, which was not included in the analysis in [14].

## A.3 Proofs for Chapter 5

### A.3.1 Derivation of Algorithm 4

Algorithm 4 builds upon the original LAMA algorithm Algorithm 2:

$$\begin{aligned}
\hat{\mathbf{z}}^t &= \hat{\mathbf{s}}^t + \mathbf{H}^{\mathrm{H}}\mathbf{r}^t \\
\hat{\mathbf{s}}^{t+1} &= \mathsf{F}(\hat{\mathbf{z}}^t, N_0 + \hat{\tau}_t^2)) \\
\hat{\tau}_{t+1}^2 &= \beta\langle \mathsf{G}(\hat{\mathbf{z}}^t, N_0 + \hat{\tau}_t^2)\rangle \\
\mathbf{r}^{t+1} &= \mathbf{y} - \mathbf{H}\hat{\mathbf{s}}^{t+1} + \tfrac{\hat{\tau}_{t+1}^2}{N_0+\hat{\tau}_t^2}\mathbf{r}^t.
\end{aligned}$$

We use the facts that $\mathbf{y}^{\mathrm{MF}} = \mathbf{H}^{\mathrm{H}}\mathbf{y}$ and $\mathbf{G} = \mathbf{H}^{\mathrm{H}}\mathbf{H}$. Then, the residual $\mathbf{r}^{t+1}$ can be simplified to:

$$\begin{aligned}
\mathbf{H}^{\mathrm{H}}\mathbf{r}^{t+1} &= \mathbf{H}^{\mathrm{H}}\mathbf{y} - \mathbf{H}^{\mathrm{H}}\mathbf{H}\hat{\mathbf{s}}^{t+1} + \tfrac{\hat{\tau}_{t+1}^2}{N_0+\hat{\tau}_t^2}\mathbf{H}^{\mathrm{H}}\mathbf{r}^t \\
&= \mathbf{y}^{\mathrm{MF}} - \mathbf{G}\hat{\mathbf{s}}^{t+1} + \tfrac{\hat{\tau}_{t+1}^2}{N_0+\hat{\tau}_t^2}\mathbf{H}^{\mathrm{H}}\mathbf{r}^t.
\end{aligned}$$

We are done as we can define $\mathbf{v}^{t+1} = \mathbf{H}^{\mathrm{H}}\mathbf{r}^t = \hat{\mathbf{z}}^t - \hat{\mathbf{s}}^t$.

### A.3.2 Proof of Lemma 28

As in (5.3), we write the estimate $\hat{z}_{c,u}$ for UE $u$ at cluster $c$ as $\hat{z}_{c,u} = s_{0,u} + e_{c,u}$, where $e_{c,u}$ represents residual interference and noise with known error variance $\mathbb{E}_{e_{c,u}}[|e_{c,u}|^2] = \sigma_{c,u}^2$. At UE $u$, optimal fusion is $\hat{z}_u = \sum_{c=1}^{C} \nu_{c,u}\hat{z}_{c,u}$ so that $\sum_{c=1}^{C}\nu_{c,u} = 1$. Hence, the fused estimate is $\hat{z}_u = s_{0,u} + \sum_{c=1}^{C}\nu_{c,u}e_{c,u}$, with the following post-fusion SINR:

$$\mathsf{SINR} = \frac{\mathbb{E}_{s_{0,u}}[|s_{0,u}|^2]}{\mathbb{E}_{\mathbf{e}_u}\left[\left|\sum_{c=1}^{C}\nu_{c,u}e_{c,u}\right|^2\right]} = \frac{E_s}{\sum_{c=1}^{C}\nu_{c,u}^2\sigma_{c,u}^2}. \tag{A.24}$$

Here, we used the assumption that the residual interference and noise terms $e_{c,u}$ are zero mean and uncorrelated across clusters $c = 1, \ldots, C$. We are now interested in maximizing the post-fusion SINR in (A.24) subject to $\sum_{c=1}^{C} \nu_{c,u} = 1$. Using the method of Lagrange multipliers, it is easy to see that

$$\nu_{c,u} = \frac{1}{\sigma_{c,u}^2} \left( \sum_{c'=1}^{C} \frac{1}{\sigma_{c',u}^2} \right)^{-1}, \quad c = 1, \ldots, C. \tag{A.25}$$

### A.3.3 Proof of Lemma 29

For Rayleigh-fading channels, each entry in the partial channel matrix $\mathbf{H}_c$ is distributed as $\mathcal{CN}(0, 1/B)$. To ensure that the expected column-norm of $\mathbf{H}_c$ is one, we normalize the per-cluster input-output relation in (2.1) by $1/\sqrt{w_c}$. This normalization amplifies the noise variance by $1/w_c$ in each cluster. In addition, since overall system dimension is $Bw_c \times U$, the resulting system ratio is given by $\beta = U/(Bw_c) = \beta/w_c$. By realizing that $N_0/w_c$ is the per-cluster noise variance, the fixed-point equation follows immediately from Theorems 26 and 2 for linear and LAMA-based equalization, respectively.

### A.3.4 Proof of Theorem 30

To simplify notation, we omit the UE index $u$. The proof follows from (5.8) in Lemma 28. The first expression in (5.9) is trivial whereas the second expression

is obtained as follows:

$$\beta \sum_{c=1}^{C} \nu_c \Psi(\bar{\sigma}_c^2) = \left( \sum_{c=1}^{C} \frac{1}{\bar{\sigma}_c^2} \right)^{-1} \sum_{c=1}^{C} \frac{\beta \Psi(\bar{\sigma}_c^2)}{\bar{\sigma}_c^2}$$
$$= \left( \sum_{c=1}^{C} \frac{1}{\bar{\sigma}_c^2} \right)^{-1} \sum_{c=1}^{C} \left( w_c - \frac{N_0}{\bar{\sigma}_c^2} \right) = \left( \sum_{c=1}^{C} \frac{1}{\bar{\sigma}_c^2} \right)^{-1} - N_0.$$

### A.3.5 Proof of Lemma 31

We first show when equality holds. The case for $C = 1$ is trivial because the PD and FD architectures are equivalent for $C = 1$. The case for $\beta \to 0$ is also straightforward because $\sigma_c^2 = \sigma_{\text{FD}}^2 = \sigma_{\text{PD}}^2 = N_0$. For MF, we have $\sigma_{\text{FD}}^2 = N_0 + \beta \sum_{c=1}^{C} \nu_c \, \mathbb{V}\text{ar}_S[S] = N_0 + \beta \, \mathbb{V}\text{ar}_S[S] = \sigma_{\text{PD}}^2$.

Let us now assume that $\beta > 0$. We show that $\sigma_c^2 > \sigma_{\text{PD}}^2$ by re-writing the fixed-point solutions as [1]: $\sigma_c^2 = \sup\{\sigma^2 : N_0 + \beta\Psi(\sigma^2) \geq w_c\sigma^2\}$ and $\sigma_{\text{PD}}^2 = \sup\{\sigma^2 : N_0 + \beta\Psi(\sigma^2) \geq \sigma^2\}$. Note that $N_0 > 0$, so both $\sigma_c^2$ and $\sigma_{\text{PD}}^2$ are strictly positive. It is easy to see that $\sigma_{\text{PD}}^2 \neq \sigma_c^2$ because $\sigma_{\text{PD}}^2 = N_0 + \beta\Psi(\sigma_{\text{PD}}^2) > w_c\sigma_{\text{PD}}^2$. Since $\Psi(\sigma^2) \to \mathbb{V}\text{ar}_S[S]$ as $\sigma^2 \to \infty$ and $\Psi(\sigma^2)$ is continuous [194], there exists a $\sigma_c^2 > \sigma_{\text{PD}}^2$ that satisfies $N_0 + \beta\Psi(\sigma_c^2) = w_c\sigma_c^2$ by the intermediate value theorem.

Finally, we use [194, Prop. 9] to see that $\Psi(\sigma^2)$ is strictly increasing for $\sigma^2 > 0$ for LAMA. For ZF and MMSE, this also holds by inspection of $\mathrm{d}\Psi(\sigma^2)/\mathrm{d}\sigma^2 > 0$. Thus, the result $\sigma_{\text{FD}}^2 > \sigma_{\text{PD}}^2$ follows directly from Lemma 28 since

$$\sigma_{\text{FD}}^2 = N_0 + \beta \sum_{c=1}^{C} \nu_c \Psi(\sigma_c^2) > N_0 + \beta \sum_{c=1}^{C} \nu_c \Psi(\sigma_{\text{PD}}^2) = \sigma_{\text{PD}}^2.$$

### A.3.6 Proof of Lemma 32

The proof is straightforward and follows from Theorem 26 and Lemma 28. Given that cluster $c$ has $Bw_c > \beta$ antennas across all clusters $C$, the input-output relation of cluster $c$ in the large-system limit under ZF equalization results in a AWGN channel with decoupled noise variance: $\sigma_c^2 = \frac{N_0}{w_c - \beta}$. The proof follows from Lemma 28 noting that $\sum_{c=1}^{C} w_c = 1$:

$$\sigma_{\text{FD}}^2 = \left( \sum_{c=1}^{C} \frac{1}{\sigma_c^2} \right)^{-1} = \left( \sum_{c=1}^{C} \frac{w_c - \beta}{N_0} \right)^{-1} = \frac{N_0}{1 - C\beta}.$$

### A.3.7 Proof of Lemma 33

The proof follows Theorem 26 with the fixed-point equation

$$w_c \sigma_c^2 = N_0 + \beta \frac{E_s}{E_s + \sigma_c^2} \sigma_c^2,$$

which results in the following $\mathsf{SINR}$ expression for cluster $C$ for L-MMSE with the FD architecture:

$$\mathsf{SINR}_{\text{FD},c}^{\text{L-MMSE}} = \frac{1}{2} \Bigg( \sqrt{ \left( 1 - \frac{E_s}{N_0} (w_c - \beta) \right)^2 + 4 \frac{E_s}{N_0} w_c } - \left( 1 - \frac{E_s}{N_0} (w_c - \beta) \right) \Bigg).$$

Lemma 33 follows from $\mathsf{SINR}_{\text{FD}}^{\text{L-MMSE}} = \sum_{c=1}^{C} \mathsf{SINR}_{\text{FD},c}^{\text{L-MMSE}}$.

### A.3.8 Proof of Lemma 34

The proof of Lemma 34 starts from (5.11). Let us denote $\overline{\mathsf{SINR}}_{\text{FD}}^{\text{L-MMSE}}$ as $\mathsf{SINR}_{\text{FD}}^{\text{L-MMSE}}$ when $w_1 = 0$ and $w_2 = \cdots = w_C = 0$. We also define $\overline{\beta} = 1 - \beta$.

Then,

$$\begin{aligned}\max_{\mathbf{w}} \mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}} &\geq \overline{\mathsf{SINR}}_{\mathrm{FD}}^{\text{L-MMSE}} \\ &= \frac{1}{2}\left(\sqrt{\left(1-\frac{E_s}{N_0}\overline{\beta}\right)^2+4\frac{E_s}{N_0}}-\left(1-\frac{E_s}{N_0}\overline{\beta}\right)\right) \\ &\quad+\frac{C-1}{2}\left(\sqrt{\left(1-\frac{E_s}{N_0}\beta\right)^2}-\left(1-\frac{E_s}{N_0}\beta\right)\right) \\ &= \frac{1}{2}\left(\sqrt{\left(1-\frac{E_s}{N_0}\overline{\beta}\right)^2+4\frac{E_s}{N_0}}-\left(1-\frac{E_s}{N_0}\overline{\beta}\right)\right) \\ &\overset{(a)}{=} \mathsf{SINR}_{\mathrm{PD}}^{\text{L-MMSE}},\end{aligned}$$

where $(a)$ follows from (L-MMSE) in Corollary 27. Since we know from Lemma 31 that $\mathsf{SINR}_{\mathrm{PD}}^{\text{L-MMSE}} \geq \mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}}$, we have that

$$\max_{\mathbf{w}} \mathsf{SINR}_{\mathrm{FD}}^{\text{L-MMSE}} = \overline{\mathsf{SINR}}_{\mathrm{FD}}^{\text{L-MMSE}} = \mathsf{SINR}_{\mathrm{PD}}^{\text{L-MMSE}}.$$

### A.3.9 Proof of Lemma 35

The proof of Lemma 35 starts from (5.11) with the definition

$$f(w,\alpha) = \sqrt{(1-\alpha(w-\beta))^2+4\alpha w}.$$

Note that $f(w,\alpha)$ is convex in $w$ as $f''(w,\alpha) \geq 0$ for $\alpha \geq 0$. The final step follows from Jensen's inequality, which implies

$$\frac{1}{C}\sum_{c=1}^{C} f(w_c,\alpha) \geq f\left(\frac{1}{C}\sum_{c=1}^{C} w_c,\alpha\right),$$

where equality holds if $w_1 = w_2 = \cdots = w_C$.

## A.4 Proof of Chapter 6

### A.4.1 Proof of Lemma 36

Since all of the possible values in the exponent of (6.4), i.e., $\ell - \ell'$, are integers ranging from $-(L-1)$ to $L-1$, the Gram matrix $\mathbf{G}_\omega$ in (6.4) is a polynomial with degree no larger than $2L-1$. Consequently, the Gram matrices $\mathbf{G}_\omega$ in (6.4) for all subcarriers $\omega = 0, \ldots, W$ are fully determined from $2L-1$ distinct and non-zero Gram-matrix base-points.

### A.4.2 Proof of Lemma 37

Evidently, (6.9) is non-negative, i.e., $f_L(\phi) \geq 0$. To show that $f_L(\phi) \leq 1$, we use the fact that the Fejér kernel $Lf_L(\phi) = L^{-1}\frac{1-\cos(L\phi)}{1-\cos(\phi)}$ is upper bounded by $L$ [170, Eq. 1.2.24]. Now, we show that $f_L(\phi)$ is monotonically decreasing in $\phi \in [0, 2\pi/L]$ for $L > 1$. We start by defining an auxiliary function $g(\phi) = \sin(L\phi/2)/\sin(\phi/2)$, so that $f_L(\phi) = g_L^2(\phi)/L^2$. For $L > 1$, $\phi \in (\pi/L, 2\pi/L]$, $\sin(L\phi/2)$ and $\sin(\phi/2)$ are monotonically decreasing and increasing respectively, and hence, $g(\phi)$ and $f_L(\phi)$, are monotonically decreasing. For $\phi \in [0, \pi/L]$, the derivative of $g(\phi)$ with respect to $\phi$ is given by:

$$\frac{\mathrm{d}g(\phi)}{\mathrm{d}\phi} = \frac{(L\cos(L\phi/2)\sin(\phi/2) - \cos(\phi/2)\sin(L\phi/2))}{2\sin^2(\phi/2)},$$

and we have that

$$\frac{L\cos(L\phi/2)\sin(\phi/2)}{\cos(\phi/2)\sin(L\phi/2)} = \frac{L\tan(\phi/2)}{\tan(L\phi/2)} < 1.$$

Hence, $g'(\phi) < 0$, and therefore, $f_L(\phi)$ is monotonically decreasing in $\phi \in [0, 2\pi/L]$.

### A.4.3 Proof of Theorem 38

Suppose we use $\mathbf{G}_{p_k}$ at base point $p_k$ to approximate $\mathbf{G}_w$ at the target subcarrier index $\omega$. Hence, the MSE in (6.8) is given by the following expression:

$$\text{0-MSE}_{\omega}^{(m,n)} = \mathbb{E}\Big[\big|[\widetilde{\mathbf{G}}_{\omega}]_{m,n} - [\mathbf{G}_{\omega}]_{m,n}\big|^2\Big]. \tag{A.26}$$

We will obtain an analytical expression for (A.26) with imperfect CSI and the BS-antenna correlation model introduced in (6.3) and (6.2), respectively. We start with expressing the channel matrix $\mathbf{H}_{\omega}$ from (6.3) as

$$\mathbf{H}_{\omega} \overset{(a)}{=} \sum_{\ell=0}^{W-1}\left(\widehat{\mathbf{H}}_{\ell}\mathbb{I}(\ell < L) + \frac{\sigma}{\sqrt{W}}\tilde{\mathbf{E}}_{\ell}\right)\exp\left(-\frac{j2\pi\omega\ell}{W}\right), \tag{A.27}$$

where $(a)$ follows from noting that the DFT is orthogonal (with normalization constant) with each entries of $\tilde{\mathbf{E}}_{\ell}$ distributed $\mathcal{CN}(0,1)$. Hence, the TD channel matrix under imperfect CSI is given as $\widehat{\mathbf{H}}_{\ell}^{\sigma} = \widehat{\mathbf{H}}_{\ell}\mathbb{I}(\ell < L) + \frac{\sigma}{\sqrt{W}}\tilde{\mathbf{E}}_{\ell}$.

Now, use the correlation model introduced in (6.2). We note that the $B \times B$ BS correlation matrix $\mathbf{R} = (1-\delta)\mathbf{I}_B + \delta\mathbf{1}_B$ can be expressed by $(\alpha\mathbf{I}_B + \beta\mathbf{1}_B)^2 = \mathbf{R}$ where $\alpha = \sqrt{1-\delta}$, and $\beta = \frac{1}{B}(\sqrt{1+(B-1)\delta} - \alpha)$. Since the matrix $(\alpha\mathbf{I}_B + \beta\mathbf{1}_B)$ is symmetric, we note that the $\widehat{\mathbf{H}}_{\ell}$ is expressed as $\widehat{\mathbf{H}}_{\ell} = (\alpha\mathbf{I}_B + \beta\mathbf{1}_B)\widehat{\mathbf{H}}_{\ell}^{\text{uncor}}$ so that

$$[\widehat{\mathbf{H}}_{\ell}]_{b,u} = \alpha[\widehat{\mathbf{H}}_{\ell}^{\text{uncor}}]_{b,u} + \beta\sum_{b=1}^{B}[\widehat{\mathbf{H}}_{\ell}^{\text{uncor}}]_{b,u} \tag{A.28}$$

For our derivation of the MSE, we will utilize the following auxiliary function:

$$R_{\ell\ell'} = \exp\left(j\frac{2\pi p_k}{W}(\ell-\ell')\right) - \exp\left(j\frac{2\pi\omega}{W}(\ell-\ell')\right). \tag{A.29}$$

By substituting $R_{\ell\ell'}$ into $[\widetilde{\mathbf{G}}_{\omega}]_{m,n} - [\mathbf{G}_{\omega}]_{m,n}$ in (A.26), we obtain the following

expression for the 0th order MSE:

$$\begin{aligned}
\text{0-MSE}_{\omega}^{(m,n)} &= \mathbb{E}\left[\left|\sum_{\ell,\ell'=0}^{W-1}\sum_{b=1}^{B}[\widehat{\mathbf{H}}_{\ell}^{\sigma}]_{b,m}^{*}[\widehat{\mathbf{H}}_{\ell'}^{\sigma}]_{b,n}R_{\ell\ell'}\right|^{2}\right] \\
&= \sum_{\ell_1,\ell_2=0}^{W-1}\sum_{\ell_3,\ell_4=0}^{W-1}\sum_{b_1,b_2=1}^{B}R_{\ell_1\ell_2}^{*}R_{\ell_3\ell_4}\mathbb{E}\left[[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_1,m}[\widehat{\mathbf{H}}_{\ell_2}^{\sigma}]_{b_1,n}^{*}[\widehat{\mathbf{H}}_{\ell_3}^{\sigma}]_{b_2,m}^{*}[\widehat{\mathbf{H}}_{\ell_4}^{\sigma}]_{b_2,n}\right] \\
&\overset{(a)}{=} \sum_{\substack{\ell_1,\ell_2=0\\ \ell_1\neq\ell_2}}^{W-1}|R_{\ell_1\ell_2}|^{2}\sum_{b_1,b_2=1}^{B}\mathbb{E}\left[[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_2,m}^{*}[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_1,m}\right]\mathbb{E}\left[[\widehat{\mathbf{H}}_{\ell_2}^{\sigma}]_{b_1,n}^{*}[\widehat{\mathbf{H}}_{\ell_2}^{\sigma}]_{b_2,n}\right]
\end{aligned} \tag{A.30}$$

where $(a)$ follows from $R_{\ell_1\ell_2}=0$ if $\ell_1=\ell_2$, and independence and the zero-mean assumption on the TD channel for $\ell_1\neq\ell_2$, which enforces $\ell_1=\ell_3$ and $\ell_2=\ell_4$ for (A.30). By inspection of (A.30), we observe that $\text{0-MSE}_{\omega}^{(m,n)}$ is independent of $m$ and $n$ and thus, the MSE of the off-diagonal and diagonal entries are equal. We simplify (A.30) for imperfect CSI and the BS-antenna correlation model. We first note that

$$\begin{aligned}
\mathbb{E}\left[[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_2,m}^{*}[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_1,m}\right] &= \frac{\sigma^2}{W}\mathbb{I}(b_1=b_2) \\
&\quad+\frac{1}{BL}(\alpha^2\mathbb{I}(b_1=b_2)+2\alpha\beta+\beta^2B)\mathbb{I}(\ell_1<L) \\
&= \frac{\sigma^2}{W}\mathbb{I}(b_1=b_2)+\frac{1}{BL}(\alpha^2\mathbb{I}(b_1=b_2)+\delta)\mathbb{I}(\ell_1<L),
\end{aligned}$$

where the last step is obtained by noting that $2\alpha\beta+\beta^2B=\delta$. Therefore, the inner sum is evaluated by

$$\begin{aligned}
&\sum_{b_1,b_2=1}^{B}\mathbb{E}\left[[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_2,m}^{*}[\widehat{\mathbf{H}}_{\ell_1}^{\sigma}]_{b_1,m}\right]\mathbb{E}\left[[\widehat{\mathbf{H}}_{\ell_2}^{\sigma}]_{b_1,n}^{*}[\widehat{\mathbf{H}}_{\ell_2}^{\sigma}]_{b_2,n}\right] \\
&= \frac{B\sigma^4}{W^2}+\frac{\sigma^2}{LW}(\alpha^2+\delta)(\mathbb{I}(\ell_1<L)+\mathbb{I}(\ell_2<L)) \\
&\quad+\frac{1}{BL^2}(\alpha^2(\alpha^2+2\delta)+B\delta^2)\mathbb{I}(\ell_1<L)\mathbb{I}(\ell_2<L).
\end{aligned} \tag{A.31}$$

Now, we simplify (A.30) using the results from (A.31) with the fact that $\alpha=$

$\sqrt{1-\delta}$ and $\sum_{\ell_1=0}^{W-1} |R_{\ell_1\ell_2}|^2 = 2W$ so that $\sum_{\ell_1,\ell_2=0}^{W-1} |R_{\ell_1\ell_2}|^2 = 2W^2$. Hence,

$$\begin{aligned}
\text{0-MSE}_{\omega}^{(m,n)} &= 2B\sigma^4 + 4\sigma^2 + \frac{1+\delta^2(B-1)}{BL^2} \sum_{\substack{\ell_1,\ell_2=0 \\ \ell_1\neq\ell_2}}^{L-1} |R_{\ell_1\ell_2}|^2 \\
&= 2B\sigma^4 + 4\sigma^2 + \frac{2(1+\delta^2(B-1))}{B}\Big(1 - \sum_{\ell=0}^{L-1}\sum_{\Delta\ell=-\ell}^{\ell} \frac{e^{j\theta\Delta\ell}}{L^2}\Big) \\
&\overset{(a)}{=} 2\sigma^2(2+B\sigma^2) + \frac{2}{B}(1+\delta^2(B-1))(1-f_L(\theta)),
\end{aligned}$$

where $(a)$ comes from the definition of Fejér kernel [170]. Note that we defined the shorthand variable $\theta = \frac{2\pi}{W}(p_k - \omega)$.

The proof can be generalized to per-UE large-scale fading by expressing the channel matrix as $\overline{\mathbf{H}}_{\omega} = \mathbf{H}_{\omega}\mathbf{D}$, where $\mathbf{H}_{\omega}$ was defined in (A.27) and the diagonal matrix $\mathbf{D}$ contains the large-scale fading coefficients for the UEs on the main diagonal. In addition, the proof can be generalized to receive-side correlation matrices $\mathbf{R}$ by rewriting $\widehat{\mathbf{H}}_{\ell}$ in (A.28) with $\widehat{\mathbf{H}}_{\ell} = \sqrt{\mathbf{R}}\widehat{\mathbf{H}}_{\ell}^{\text{uncor}}$. A corresponding analysis is left for future work.

### A.4.4 Proof of Corollary 40

Since $f_L(\phi)$ is non-negative, it is obvious that $\max_{\omega\in\Omega}\{\text{0-MSE}_{\omega}\} \leq \varepsilon_{\text{CSI}} + 2(1+\varepsilon_{\text{cor}})/B$ for all $\omega \in \Omega$. The equality is satisfied if $(p-\omega)/W = b/L$ for some integer $b > 0$ so that $f_L\big(\frac{2\pi}{W}(p-\omega)\big) = f_L\big(\frac{b2\pi}{L}\big) = 0$. We note that this can only happen if $d_{\max} \geq W/L$, where $d_{\max}$ is the maximum distance between any target subcarrier $\omega$ point and its nearest base point; this is due to the fact that $\frac{2\pi}{W}(p-\omega) \leq \frac{2\pi}{W}d_{\max} < \frac{2\pi}{L}$.

Assume $d_{\max} < W/L$. Then, by Lemma 37, the maximum MSE of 0th order

interpolation is given by:

$$\begin{aligned}\max_{\omega\in\Omega}\{\text{0-MSE}_\omega\} &= \varepsilon_{\text{CSI}} + \frac{2}{B}(1+\varepsilon_{\text{cor}})\left(1 - \min_{\omega\in\Omega\setminus\mathcal{P}} f_L\left(\frac{2\pi}{W}(p_k-\omega)\right)\right)\\ &= \varepsilon_{\text{CSI}} + \frac{2}{B}(1+\varepsilon_{\text{cor}})\left(1 - f_L\left(\frac{2\pi}{W}d_{\max}\right)\right).\end{aligned}$$

### A.4.5 Proof of Theorem 41

The proof is similar to that of Theorem 38 in Appendix A.4.3. We start by defining the following auxiliary function

$$Q_{\ell\ell'} = \lambda_\omega \exp\left(j\frac{2\pi p_k}{W}\Delta\ell\right) + (1-\lambda_\omega)\exp\left(j\frac{2\pi p_{k+1}}{W}\Delta\ell\right) - \exp\left(j\frac{2\pi\omega}{W}\Delta\ell\right),$$

where we introduced the variable $\Delta\ell = \ell - \ell'$. The result is obtained by substituting $Q_{\ell\ell'}$ in place of $R_{\ell\ell'}$ at (A.30) in Appendix A.4.3. Note that $\sum_{\ell_1=0}^{W-1}|Q_{\ell_1\ell_2}|^2 = 2(1-\lambda_\omega(1-\lambda_\omega))W$ which shows that $\sum_{\ell_1,\ell_2=0}^{W-1}|Q_{\ell_1\ell_2}|^2 = 2(1-\lambda_\omega(1-\lambda_\omega))W^2$.

### A.4.6 Proof of Corollary 42

Without loss of generality, we will assume that the target subcarrier index $\omega$ is closer to $p_{k+1}$ so that $\lambda_\omega \in [0, 0.5]$. We will assume that $p_k < \omega < p_{k+1}$ so that $d_k = p_{k+1} - p_k > 0$. Using the results from Appendix A.4.3 and Appendix A.4.5, the difference of $\text{1-MSE}_\omega$ and $\text{0-MSE}_\omega$ is given by:

$$\text{1-MSE}_\omega - \text{0-MSE}_\omega = -\varepsilon_{\text{CSI}}\lambda_\omega(1-\lambda_\omega)) + \frac{1+\varepsilon_{\text{cor}}}{BL^2}\sum_{\ell_1,\ell_2=0}^{L-1}(|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2).$$

Without loss of generality, we assume that $\ell_1 - \ell_2 > 0$ since $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2$ is even and is 0 if $\ell_1 = \ell_2$. We simplify the term $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2$ by denoting

$\phi = \theta(\ell_1 - \ell_2) > \theta$, where $\theta = \frac{2\pi}{W} d_k$ and expand the expression $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2$ as follows:

$$\begin{aligned} |Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2 &\overset{(a)}{=} 2\lambda_\omega\big((1-\lambda_\omega)(\cos(\phi)-1) \\ &\quad - (\cos((1-\lambda_\omega)\phi) - \cos(\lambda_\omega\phi))\big) \\ &\overset{(b)}{=} 2\lambda_\omega \sin(\phi/2)(\sin((1-2\lambda_\omega)\phi/2) - (1-\lambda_\omega)\sin(\phi/2)). \end{aligned} \tag{A.32}$$

Here, $(a)$ follows from the definition of $Q_{\ell_1\ell_2}$ and $R_{\ell_1\ell_2}$ and $(b)$ is a results from simplifying the expression $\cos((1-\lambda_\omega)\phi) - \cos(\lambda_\omega\phi)) = -2\sin((1-2\lambda_\omega)\phi/2)\sin(\phi/2)$.

We first note that since $\varepsilon_{\text{CSI}} = 2\sigma^2(2 + B\sigma^2) \geq 0$ and by (A.32), 1-$\text{MSE}_\omega$ = 0-$\text{MSE}_\omega$ if $\varepsilon_{\text{CSI}} = 0$, and $L = 1$ or $\lambda_\omega = 0$. This behavior can be explained intuitively because when $\varepsilon_{\text{CSI}} = 0$ and $L = 1$, then the channel is flat across all subcarriers, and hence 1-$\text{MSE}_\omega$ = 0-$\text{MSE}_\omega$

Hence, we now show that 1-$\text{MSE}_\omega$ < 0-$\text{MSE}_\omega$ for $L > 1$ and $\lambda_\omega \neq 0$ by showing that $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2 < 0$. First note that $\ell_1 - \ell_2 < L$ and if $d_k < W/(3L)$, then $\theta < \phi < L\theta = \frac{2\pi}{W} L d_k < \frac{2\pi}{3}$. Hence, $\sin(\phi/2) > 0$. Therefore, showing that (A.32) is negative is equivalent to:

$$g(\lambda_\omega) = \frac{\sin((1-2\lambda_\omega)\phi/2)}{1-\lambda_\omega} < \sin(\phi/2). \tag{A.33}$$

We now prove (A.33) by noting that $g(\lambda_\omega) = \sin(\phi/2)$ if $\lambda_\omega = 0$ and $g'(\lambda_\omega) < 0$ for all $\lambda_\omega \in (0, 0.5]$ so $g(\lambda_\omega)$ is monotonically decreasing in $(0, 0.5]$. The proof is straightforward by:

$$g'(\lambda_\omega) = \frac{-\phi\cos((1-2\lambda_\omega)\phi/2)}{1-\lambda_\omega} + \frac{\sin((1-2\lambda_\omega)\phi/2)}{(1-\lambda_\omega)^2}, \tag{A.34}$$

and, hence, $g'(\lambda_\omega) < 0$ in (A.34) can be expressed as:

$$\tan((1-2\lambda_\omega)\phi/2) < (1-\lambda_\omega)\phi. \tag{A.35}$$

To show (A.35), we introduce the shorthand notation $\gamma = 1 - 2\lambda_\omega \in [0, 1)$. With the new notation $\gamma$, the proof is straightforward by:

$$\begin{aligned}\tan((1-2\lambda_\omega)\phi/2) &\overset{(a)}{\leq} \gamma \tan(\phi/2) \overset{(b)}{\leq} \gamma\phi < (\gamma + \lambda_\omega)\phi \\ &= (1-\lambda_\omega)\phi,\end{aligned}$$

where $(a)$ follows from the convexity of $\tan(x)$ in $x \in [0, \pi/2)$ and $(b)$ follows from $\tan(\phi/2) < \phi$ for all $\phi \in (0, 2\pi/3]$. Since $g'(\lambda_\omega) < 0$ for all $\lambda_\omega \in (0, 0.5]$, $g(\lambda_\omega)$ is monotonically decreasing and thus, from (A.32), it follows that $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2 < 0$ for $L > 1$.

We conclude by noting that a sharper upper bound on $d_k$ can be obtained by directly computing the bounds for $\mathsf{1\text{-}MSE}_\omega - \mathsf{0\text{-}MSE}_\omega$, i.e.,

$$\begin{aligned}\mathsf{1\text{-}MSE}_\omega - \mathsf{0\text{-}MSE}_\omega &= -\varepsilon_{\text{CSI}}\lambda_\omega(1-\lambda_\omega)) + \frac{2\lambda_\omega(1+\varepsilon_{\text{cor}})}{B} \\ &\times \Big(f_L(\lambda_\omega\theta) - f_L((1-\lambda_\omega)\theta) - (1-\lambda_\omega)(1-f_L(\theta))\Big) < 0,\end{aligned}$$

for all $\lambda_\omega \in (0, 0.5]$, but we leave an analysis of such refined bounds for future work.

## A.5 Proofs of Chapter 7

### A.5.1 Proof of Theorem 43

The proof is easy to see as

$$\max_{a \in \mathcal{X}_j^1} \sum_{j=1}^{Q/2} \frac{2[a]_j - 1}{2} \Lambda_{u,j}^{\mathsf{a}} = \frac{1}{2}\Lambda_{u,j}^{\mathsf{a}} + \frac{1}{2}\sum_{\substack{k=1 \\ k \neq j}}^{Q/2} \left|\Lambda_{u,k}^{\mathsf{a}}\right|,$$

as there exists a symbol $a \in \mathcal{X}_j^1$ where $[a]_j = 0$ if $\Lambda^{\mathrm{a}}_{u,j} < 0$, and $[a]_j = 1$ otherwise. Therefore,

$$\begin{aligned}\lambda^{\text{prior}}_{u,j} &= \max_{a \in \mathcal{X}_j^1} \sum_{j=1}^{Q/2} \frac{2[a]_j - 1}{2} \Lambda^{\mathrm{a}}_{u,j} - \max_{a \in \mathcal{X}_j^0} \frac{1}{2} \sum_{j=1}^{Q/2} \frac{2[a]_j - 1}{2} \Lambda^{\mathrm{a}}_{u,j} \\ &= \frac{1}{2}\Lambda^{\mathrm{a}}_{u,j} + \frac{1}{2} \sum_{\substack{k=1 \\ k \neq j}}^{Q/2} \left|\Lambda^{\mathrm{a}}_{u,k}\right| + \frac{1}{2}\Lambda^{\mathrm{a}}_{u,j} - \frac{1}{2} \sum_{\substack{k=1 \\ k \neq j}}^{Q/2} \left|\Lambda^{\mathrm{a}}_{u,k}\right| = \Lambda^{\mathrm{a}}_{u,j}.\end{aligned}$$

## BIBLIOGRAPHY


[1] A. Maleki, "Approximate message passing algorithms for compressed sensing," Ph.D. dissertation, Stanford University, Jan. 2011.

[2] B. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, March 2003.

[3] C. Studer, A. Burg, and H. Bölcskei, "Soft-output sphere decoding: Algorithms and VLSI implementation," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 290–300, Feb. 2008.

[4] C. Studer and H. Bölcskei, "Soft–input soft–output single tree-search sphere decoding," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4827–4842, Oct. 2010.

[5] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[6] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *Proc. Allerton Conf. Commun., Contr., Comput.*, Sept. 2011, pp. 545–550.

[7] K. Alnajjar, P. Smith, and G. Woodward, "Low complexity V-BLAST for massive MIMO," in *Proc. Aus. Commun. Theory Workshop (AusCTW)*, Feb. 2014, pp. 22–26.

[8] Z. Wu, C. Zhang, Y. Xue, S. Xu, and X. You, "Efficient architecture for soft-output massive MIMO detection with Gauss-Seidel method," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2016, pp. 1886–1889.

[9] A. Maleki, L. Anitori, Z. Yang, and R. Baraniuk, "Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP)," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4290–4308, Jul. 2013.

[10] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[11] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.

[12] D. Guo and C.-C. Wang, "Asymptotic mean-square optimality of belief propagation for sparse linear systems," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2006, pp. 194–198.

[13] D. Tse and S. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[14] C. Thrampoulidis, W. Xu, and B. Hassibi, "Symbol error rate performance of box-relaxation decoders in massive MIMO," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3377–3392, Jul. 2018.

[15] C. Thrampoulidis, E. Abbasi, W. Xu, and B. Hassibi, "BER analysis of the box relaxation for BPSK signal recovery," in *Proc. IEEE Int'l Conf. Acoust., Speech and Signal Process. (ICASSP)*, Mar. 2016, pp. 3776–3780.

[16] A. Puglielli, N. Narevsky, P. Lu, T. Courtade, G. Wright, B. Nikolic, and E. Alon, "A scalable massive MIMO array architecture based on common modules," in *IEEE Intl. Conf. Commun. Workshop (ICCW)*, June 2015, pp. 1310–1315.

[17] K. Li, R. R. Sharan, Y. Chen, T. Goldstein, J. R. Cavallaro, and C. Studer, "Decentralized baseband processing for massive MU-MIMO systems," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 4, pp. 491–507, Nov. 2017.

[18] L. Van der Perre, L. Liu, and E. G. Larsson, "Efficient DSP and circuit architectures for massive MIMO: State-of-the-art and future directions," *arXiv preprint: 1807.05882*, July 2018.

[19] S. Jacobsson, Y. Ettefagh, G. Durisi, and C. Studer, "All-digital massive MIMO with a fronthaul constraint," in *Proc. IEEE Workshop Stat. Signal Process. (SSP)*, Jun. 2018.

[20] http://www.cpri.info, *Common public radio interface*. [Online]. Available: http://www.cpri.info/

[21] A. Puglielli, A. Townley, G. LaCaille, V. Milovanović, P. Lu, K. Trotskovsky, A. Whitcombe, N. Narevsky, G. Wright, T. Courtade *et al.*, "Design of

energy-and cost-efficient massive MIMO arrays," *Proc. IEEE*, vol. 104, no. 3, pp. 586–606, Dec. 2016.

[22] M. Wu, B. Yin, G. Wang, C. Dick, J. Cavallaro, and C. Studer, "Large-scale MIMO detection for 3GPP LTE: Algorithm and FPGA implementation," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 916–929, Oct. 2014.

[23] B. Yin, "Low complexity detection and precoding for massive MIMO systems: Algorithm, architecture, and application," Ph.D. dissertation, Rice University, Dec. 2014.

[24] D. Cescato, M. Borgmann, H. Bölcskei, J. Hansen, and A. Burg, "Interpolation-based QR decomposition in MIMO-OFDM systems," in *Proc. IEEE Int'l Workshop Signal Process. Advances for Wireless Commun. (SPAWC)*, June 2005, pp. 945–949.

[25] M. Borgmann and H. Bölcskei, "Interpolation-based efficient matrix inversion for MIMO-OFDM receivers," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Nov. 2004, pp. 1941–1947 Vol.2.

[26] D. Cescato and H. Bölcskei, "Algorithms for interpolation-based QR decomposition in MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1719–1733, Apr. 2011.

[27] ——, "QR decomposition of Laurent polynomial matrices sampled on the unit circle," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4754–4761, Sep. 2010.

[28] M. Čirkić and E. Larsson, "SUMIS: Near-optimal soft-in soft-out MIMO detection with low and fixed complexity," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3084–3097, Jun. 2014.

[29] J. W. Choi, B. Lee, B. Shim, and I. Kang, "Low complexity detection and precoding for massive MIMO systems," in *Proc. IEEE Wireless Commun. & Netw. Conf. (WCNC)*, April 2013, pp. 2857–2861.

[30] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[31] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer, "FPGA design of a coordinate descent data detector for large-scale MU-MIMO," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2016, pp. 1894–1897.

[32] G. Peng, L. Liu, S. Zhou, S. Yin, and S. Wei, "A 1.58 Gbps/W 0.40 Gbps/mm$^2$ ASIC Implementation of MMSE Detection for 128 $\times$ 8 64-QAM Massive MIMO in 65 nm CMOS," *IEEE Trans. Circuits Syst. I*, vol. 65, no. 5, pp. 1717–1730, May 2018.

[33] B. Yin, M. Wu, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, "A 3.8gb/s large-scale MIMO detector for 3GPP LTE-Advanced," in *Proc. IEEE Int'l Conf. Acoust., Speech and Signal Process. (ICASSP)*, May 2014, pp. 3879–3883.

[34] H. Prabhu, J. N. Rodrigues, L. Liu, and O. Edfors, "3.6 A 60pJ/b 300Mb/s 128 $\times$ 8 Massive MIMO precoder-detector in 28nm FD-SOI," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2017, pp. 60–61.

[35] W. Tang, C. Chen, and Z. Zhang, "A 0.58mm$^2$ 2.76Gb/s 79.8pJ/b 256-QAM massive MIMO message-passing detector," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Jun. 2016, pp. 1–2.

[36] Y. Chen, C. Cheng, T. Tsai, W. Sun, Y. Ueng, and C. Yang, "A 501mW 7.6lGb/s integrated message-passing detector and decoder for polar-coded massive MIMO systems," in *VLSI Symp.*, Jun. 2017, pp. C330–C331.

[37] W. Tang, H. Prabhu, L. Liu, V. Öwall, and Z. Zhang, "A 1.8Gb/s 70.6pJ/b 128 $\times$ 16 link-adaptive near-optimal massive MIMO detector in 28nm UTBB-FDSOI," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2018, pp. 224–226.

[38] C. Jeon, R. Ghods, A. Maleki, and C. Studer, "Optimal data detection in large MIMO," *arXiv:1811.01917 [cs.IT]*, 2018.

[39] ——, "Optimality of large MIMO detection via approximate message passing," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 1227–1231.

[40] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *Int. ITG Workshop on Smart Antennas (WSA)*, Feb. 2010, pp. 189–196.

[41] R. Ghods, C. Jeon, G. Mirza, A. Maleki, and C. Studer, "Optimally-tuned nonparametric linear equalization for massive MU-MIMO systems," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2118–2122.

[42] L. Cannon, "A cellular computer to implement the Kalman filter algorithm," Ph.D. dissertation, Montana State University, USA, 1969.

[43] Ericsson, "Ericsson mobility report q2 2018 update," Tech. Rep., Jun. 2018. [Online]. Available: https://www.ericsson.com/en/mobility-report

[44] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Survey Tut.*, vol. 17, no. 4, pp. 2347–2376, 4Q 2015.

[45] TechInsights, "Samsung Galaxy S8 (SM-G950W) Teardown," Tech. Rep., Apr. 2017. [Online]. Available: https://techinsights.com/about-techinsights/overview/blog/samsung-galaxy-s8-teardown

[46] W. Lam, "The RF front-end: Unsung Hero of the Premium Smartphone," IHS Markit Technology, Tech. Rep., Jul. 2017. [Online]. Available: https://technology.ihs.com/593647/the-rf-front-end-unsung-hero-of-the-premium-smartphone

[47] *IEEE Draft Standard; Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications; Amendment 4: Enhancements for Higher Throughput*, P802.11n/D3.0, Sep. 2007.

[48] A. Bleicher, "4G gets real," *IEEE Spectr.*, vol. 51, no. 1, pp. 38–62, Jan. 2014.

[49] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, "LTE release 12 and beyond," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 154–160, Jul. 2013.

[50] F. Khan, *LTE for 4G Mobile Broadband: Air Interface Technologies and Performance*. Cambridge University Press, 2009.

[51] B. Bellalta, "IEEE 802.11ax: High-efficiency WLANs," *IEEE Wireless Commun. Mag.*, vol. 23, no. 1, pp. 38–46, Feb. 2016.

[52] S. Cherry, "Edholm's law of bandwidth," *IEEE Spectr.*, vol. 41, no. 7, pp. 58–60, Jul. 2004.

[53] T. L. Marzetta, "Non-cooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Comm.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[54] V. Morgenshtern and H. Bölcskei, "Crystallization in large wireless networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3319–3349, Oct. 2007.

[55] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.

[56] T. L. Marzetta, "How much training is required for multi-user MIMO?" in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Oct. 2006, pp. 359–363.

[57] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge Univ. Press, 2003.

[58] S. Verdú, *Multiuser Detection*. Cambridge University Press, 1998.

[59] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," *arXiv:1010.5141*, 2010.

[60] N. Lashkarian, J. Shi, and M. Forbes, "A direct learning adaptive scheme for power-amplifier linearization based on Wirtinger calculus," *IEEE Trans. Circuits Syst. I*, vol. 61, no. 12, pp. 3496–3505, Dec. 2014.

[61] R. Hunger, "An introduction to complex differentials and complex differentiability," Technische Universität München, Associate Institute for Signal Processing, Technical Report TUM-LNS-TR-07-06, 2007.

[62] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[63] D. Guo and S. Verdú, "Multiuser detection and statistical mechanics," in *Commun., Inf. and Netw. Security*. Springer, 2003, pp. 229–277.

[64] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[65] M. Wu, B. Yin, A. Vosoughi, C. Studer, J. Cavallaro, and C. Dick, "Approximate matrix inversion for high-throughput data detection in the large-scale MIMO uplink," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2013, pp. 2155–2158.

[66] W. C. Y. Lee, "Overview of cellular CDMA," *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 291–302, May 1991.

[67] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*. Addison Wesley Longman Publishing Co., Inc., 1995.

[68] K. Gilhousen, I. Jacobs, R. Padovani, A. Viterbi, J. Weaver, L.A., and I. Wheatley, C.E., "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 303–312, May 1991.

[69] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, vol. 35, no. 12, pp. 126–133, Dec 1997.

[70] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.

[71] Y. Kabashima, "A CDMA multiuser detection algorithm on the basis of belief propagation," *J. Phys. A: Math. Gen.*, vol. 36, no. 43, pp. 11 111–11 121, Oct. 2003.

[72] A. B. Brown, "Linear diophantine equations," *Math. Mag.*, vol. 31, no. 4, pp. 215–220, 1958.

[73] E. Contejean and H. Devie, "An efficient incremental algorithm for solving systems of linear diophantine equations," *Inform. Comput.*, vol. 113, no. 1, pp. 143–172, Aug. 1994.

[74] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comp.*, vol. 44, no. 170, pp. 463–471, Apr. 1985.

[75] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[76] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[77] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond*. Singapore: World Scientific, 1987.

[78] A. Montanari and D. Tse, "Analysis of belief propagation for non-linear problems: The example of CDMA (or: How to prove Tanaka's formula)," in *Proc. IEEE Inf. Theory Workshop*, March 2006, pp. 160–164.

[79] D. Guo and C.-C. Wang, "Random sparse linear systems observed via arbitrary channels: A decoupling principle," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, 2007, pp. 946–950.

[80] G. Caire, R. Müller, and T. Tanaka, "Iterative multiuser joint decoding: optimal power allocation and low-complexity implementation," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1950–1973, Sept. 2004.

[81] D. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Natl. Academy of Sciences (PNAS)*, vol. 106, no. 45, pp. 18 914–18 919, Sept. 2009.

[82] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.

[83] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 1289–1306, Jan. 2006.

[84] E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[85] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "From Denoising to Compressed Sensing," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 5117–5144, Sep. 2016.

[86] A. Montanari, *Graphical models concepts in compressed sensing, Compressed Sensing (Y.C. Eldar and G. Kutyniok, eds.)*. Cambridge University Press, 2012.

[87] D. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. Motivation and construction," in *Proc. IEEE Inf. Theory Workshop*, Jan. 2010, pp. 1–5.

[88] ——, "Message passing algorithms for compressed sensing: II. Analysis and validation," in *Proc. IEEE Inf. Theory Workshop*, Jan. 2010, pp. 1–5.

[89] A. Javanmard and A. Montanari, "State evolution for general approximate

message passing algorithms, with applications to spatial coupling," *J. Inf. Inference*, vol. 2, no. 2, pp. 115–144, Oct. 2013.

[90] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1588–1592.

[91] J. Vila and P. Schniter, "Expectation-maximization gaussian-mixture approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4658–4672, Oct. 2013.

[92] U. Kamilov, S. Rangan, A. Fletcher, and M. Unser, "Approximate message passing with consistent parameter estimation and applications to sparse learning," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2969–2985, May 2014.

[93] S. Rangan, P. Schniter, E. Riegler, A. Fletcher, and V. Cevher, "Fixed points of generalized approximate message passing with arbitrary matrices," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2013, pp. 664–668.

[94] J. Vila and P. Schniter, "Expectation-maximization bernoulli-gaussian approximate message passing," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2011, pp. 799–803.

[95] U. Kamilov, V. Goyal, and S. Rangan, "Message-passing de-quantization with applications to compressed sensing," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6270–6281, Dec. 2012.

[96] P. Maechler, C. Studer, D. E. Bellasi, A. Maleki, A. Burg, N. Felber, H. Kaeslin, and R. G. Baraniuk, "VLSI design of approximate message passing for signal restoration and compressive sensing," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 579–590, Sep. 2012.

[97] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a markov-tree prior," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3439–3448, Jul. 2012.

[98] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," in *Proc. Allerton Conf. Commun., Contr., Comput.*, Oct. 2012, pp. 815–822.

[99] Y. Ma, J. Zhu, and D. Baron, "Compressed Sensing via Universal Denoising and Approximate Message Passing," *arXiv:1407.1944*, Jul. 2014.

[100] M. Nassar, P. Schniter, and B. Evans, "A factor graph approach to joint OFDM channel estimation and decoding in impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1576–1589, Mar. 2014.

[101] M. Nabaee and F. Labeau, "Bayesian quantized network coding via generalized approximate message passing," in *Proc. Wireless Telecomunn. Symp. (WTS)*, Apr. 2014, pp. 1–7.

[102] P. Schniter, "A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels," *Proc. Intl. Workshop IEEE Sig. Proc. Adv. Wireless Commun. (SPAWC)*, pp. 246–250, Jun. 2011.

[103] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 902–915, Oct. 2014.

[104] D. Guo and S. Verdú, "Replica analysis of large-system CDMA," *Proc. IEEE Inf. Theory Workshop*, pp. 22–25, Mar. 2003.

[105] M. Mezard and A. Montanari, *Information, Physics, and Computation*. Oxford University Press, Inc., 2009.

[106] M. Bayati and A. Montanari, "The LASSO risk for Gaussian matrices," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 1997–2017, Apr. 2012.

[107] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[108] Y. Wu and S. Verdú, "MMSE dimension," *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, pp. 1463–1467, Jun. 2010.

[109] J. Jalden and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[110] D. Seethaler, J. Jaldén, C. Studer, and H. Bölcskei, "On the complexity distribution of sphere decoding," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5754–5768, Sept. 2011.

[111] T. Narasimhan and A. Chockalingam, "Channel hardening-exploiting

message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 847–860, Oct. 2014.

[112] P. Suthisopapan, K. Kasai, V. Imtawil, and A. Meesomboon, "Approaching capacity of large MIMO systems by non-binary LDPC codes and MMSE detection," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 1712–1716.

[113] C.-C. Wang and D. Guo, "Belief propagation is asymptotically equivalent to MAP detection for sparse linear systems," in *Proc. Allerton Conf. Commun., Contr., Comput.*, 2006, pp. 926–935.

[114] L. Zheng, A. Maleki, X. Wang, and T. Long, "Does $\ell_p$-minimization outperform $\ell_1$-minimization?" *arXiv:1501.03704*, Jan. 2015.

[115] G. Reeves and H. D. Pfister, "The replica-symmetric prediction for compressed sensing with Gaussian matrices is exact," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 665–669.

[116] D. L. Donoho and J. Tanner, "Counting the faces of randomly-projected hypercubes and orthants, with applications," *Discrete Comput. Geometry*, vol. 43, no. 3, pp. 522–541, Apr. 2010.

[117] ——, "Precise undersampling theorems," *Proc. IEEE*, vol. 98, no. 6, pp. 913–924, 2010.

[118] T. Tanaka, "Analysis of bit error probability of direct-sequence CDMA multiuser demodulators," in *NIPS*, 2000, pp. 315–321.

[119] D. Pirjol, "The logistic-normal integral and its generalizations," *J. Comput. Appl. Math.*, vol. 237, no. 1, pp. 460–469, Jan. 2013.

[120] G. E. Crooks, "Logistic approximation to the logistic-normal integral," Lawrence Berkeley National Laboratory, Tech. Rep. 002v4, 2013.

[121] E. Demidenko, *Generalized Linear Mixed Models*. John Wiley & Sons, Inc., 2005, pp. 329–430.

[122] A. Kuriya and T. Tanaka, "Performance degradation of AMP for small-sized problems," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, vol. Jun., 2015, pp. 2802–2806.

[123] ——, "Effects of the approximations from BP to AMP for small-sized problems," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 770–774.

[124] S. Rangan, P. Schniter, and A. Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, June 2014, pp. 236–240.

[125] O. Mangasarian and B. Recht, "Probability of unique integer solution to a system of linear equations," *Eur. J. Oper. Res.*, vol. 214, no. 1, pp. 27–30, Oct. 2011.

[126] C. Studer, T. Goldstein, W. Yin, and R. G. Baraniuk, "Democratic representations," *arXiv:1401.3420*, Apr. 2015.

[127] C. Thrampoulidis, E. Abbasi, and B. Hassibi, "Precise error analysis of regularized $M$-estimators in high dimensions," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5592–5628, Aug. 2018.

[128] S. Shamai and S. Verdu, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[129] S. Verdú, "Mismatched estimation and relative entropy," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3712–3720, Aug. 2010.

[130] Y. Ma, D. Baron, and A. Beirami, "Mismatched estimation in large linear systems," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 760–764.

[131] P. H. Tan, L. K. Rasmussen, and T. J. Lim, "Constrained maximum-likelihood detection in CDMA," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 142–153, Jan. 2001.

[132] A. Yener, R. D. Yates, and S. Ulukus, "CDMA multiuser detection: A nonlinear programming approach," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1016–1024, Jun. 2002.

[133] J. Pan, W.-K. Ma, and J. Jalden, "MIMO detection by Lagrangian dual maximum-likelihood relaxation: Reinterpreting regularized lattice decoding," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 511–524, Nov. 2014.

[134] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms

for compressed sensing," *Proc. Natl. Acad. Sci. USA*, vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

[135] A. Mousavi, A. Maleki, and R. G. Baraniuk, "Parameterless Optimal Approximate Message Passing," *arXiv:1311.0035 [cs.IT]*, Oct. 2013.

[136] ——, "Consistent Parameter Estimation for LASSO and Approximate Message Passing," *arXiv:1511.01017 [math.ST]*, Nov. 2015.

[137] Y. Eldar and A. Chan, "On the asymptotic performance of the decorrelator," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2309–2313, Sep. 2003.

[138] D. Bellasi, L. Bettini, T. Burger, C. Benkeser, Q. Huang, and C. Studer, "Compressive sensing spectrum recovery from quantized measurements in 28nm soi cmos," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, Sep. 2014, pp. 256–260.

[139] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, 2011.

[140] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE J. Sel. Topics in Sig. Proc.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[141] S. Malkowsky, J. Vieira, K. Nieman, N. Kundargi, I. Wong, V. Öwall, O. Edfors, F. Tufvesson, and L. Liu, "Implementation of low-latency signal processing and data shuffling for TDD massive MIMO systems," in *IEEE Intl. Workshop Signal Process. Syst.*, Oct. 2016, pp. 260–265.

[142] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for Massive MIMO," in *2014 IEEE Globecom Workshops*, Dec. 2014, pp. 287–293.

[143] S. Malkowsky, J. Vieira, K. Nieman, N. Kundargi, I. Wong, V. Owall, O. Edfors, F. Tufvesson, and L. Liu, "Implementation of low-latency signal processing and data shuffling for TDD massive MIMO systems," in *IEEE Workshop Sig. Proc. Systems*, Oct. 2016.

[144] Q. Yang, X. Li, H. Yao, J. Fang, K. Tan, W. Hu, J. Zhang, and Y. Zhang, "BigStation: Enabling scalable real-time signal processing in large MU-MIMO systems," in *Proc. ACM SIGCOMM Conf.*, 2013, pp. 399–410.

[145] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc. Ann. Intl. Conf. Mobile Comput. Netw. (MobiCom)*, 2012, pp. 53–64.

[146] K. Li, R. Sharan, Y. Chen, J. R. Cavallaro, T. Goldstein, and C. Studer, "Decentralized beamforming for massive MU-MIMO on a GPU cluster," in *Global Conf. Sig. Inform. Process. (GlobalSIP)*, Dec. 2016, pp. 590–594.

[147] K. Li, Y. Chen, R. Sharan, T. Goldstein, J. R. Cavallaro, and C. Studer, "Decentralized data detection for massive MU-MIMO on a Xeon Phi cluster," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2016, pp. 468–472.

[148] E. Bertilsson, O. Gustafsson, and E. G. Larsson, "A scalable architecture for massive MIMO base stations using distributed processing," in *50th Asilomar Conf. Sig., Systems, and Computers*, Nov. 2016, pp. 864–868.

[149] C. Jeon, K. Li, J. R. Cavallaro, and C. Studer, "On the achievable rates of decentralized equalization in massive MU-MIMO systems," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1102–1106.

[150] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[151] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 29, no. 2, pp. 6–14, Mar. 2015.

[152] R. Ghods, C. Jeon, G. Mirza, A. Maleki, and C. Studer, "Optimally-tuned nonparametric linear equalization for massive MU-MIMO systems," in *IEEE Intl. Symposium on Info. Theory (ISIT)*, June 2017, pp. 2118–2122.

[153] C. Jeon, A. Maleki, and C. Studer, "On the performance of mismatched data detection in large MIMO systems," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 180–184.

[154] J. Zhu, A. Beirami, and D. Baron, "Performance trade-offs in multi-processor approximate message passing," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 680–684.

[155] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[156] E. J. Candès, J. Romberg, and T.Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, Jan. 2006.

[157] J. Vila, P. Schniter, S. Rangan, F. Krzakala, and L. Zdeborova, "Adaptive damping and mean removal for the generalized approximate message passing algorithm," in *IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2015, pp. 2021–2025.

[158] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "Conjugate gradient-based soft-output detection and precoding in massive MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 3696–3701.

[159] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," 3rd Generation Partnership Project (3GPP), TS 36.211, Jan. 2016. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36211.htm

[160] L. W. Chai, P. L. Chiu, and Y. H. Huang, "Reduced-complexity interpolation-based QR decomposition using partial layer mapping," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2011, pp. 2381–2384.

[161] P. L. Chiu, L. Z. Huang, L. W. Chai, and Y. H. Huang, "Interpolation-based QR decomposition and channel estimation processor for MIMO-OFDM system," *IEEE Trans. Circuits Syst. I*, vol. 58, no. 5, pp. 1129–1141, May 2011.

[162] 3GPP, "Study on new radio (NR) access technology," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.912, Jul. 2018, version 15.0.0. [Online]. Available: http://www.3gpp.org/DynaReport/38912.htm

[163] J. Jeon, "NR wide bandwidth operations," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 42–46, Mar. 2018.

[164] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 239–241, Jun. 2002.

[165] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath,

"Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[166] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[167] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *arXiv:1111.3752v1*, 2012.

[168] C. Desset, E. Björnson, S. Kashyap, E. G. Larsson, C. Mollén, L. Liu, S. Malkowsky, H. Prabhu, Y. Liu, J. Vieira, O. Edfors, E. Karipidis, F. Dielacher, and D. V. Pop, "Distributed and centralized baseband processing algorithms, architectures, and platforms," MAMMOET, Tech. Rep. ICT-619086-D3.2, 2016.

[169] H. Bölcskei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.

[170] P. L. Butzer and R. J. Nessel, *Fourier analysis and approximation*. Academic Press New York, 1971.

[171] M. Wu, B. Yin, K. Li, C. Dick, J. R. Cavallaro, and C. Studer, "Implicit vs. explicit approximate matrix inversion for wideband massive MU-MIMO data detection," *J. Signal Process. Syst.*, vol. 90, no. 10, pp. 1311–1328, Oct. 2018.

[172] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[173] O. Castañeda, T. Goldstein, and C. Studer, "Data detection in large multi-antenna wireless systems via approximate semidefinite relaxation," *IEEE Trans. Circuits Syst. I*, vol. 63, no. 12, pp. 2334–2346, Dec. 2016.

[174] J. Céspedes, P. M. Olmos, M. Sánchez-Fernández, and F. Perez-Cruz, "Expectation propagation detection for high-order high-dimensional MIMO systems," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2840–2849, Aug. 2014.

[175] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-

correcting coding and decoding: Turbo-codes (1)," in *Proc. IEEE Int'l Conf. Commun. (ICC)*, vol. 2, May 1993, pp. 1064–1070 vol.2.

[176] J. Kim and I. Park, "A unified parallel radix-4 turbo decoder for mobile WiMAX and 3GPP-LTE," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, Sep. 2009, pp. 487–490.

[177] S. Wang, Y. Li, Z. Gao, and J. Wang, "Weighted-damped approximate message passing for compressed sensing," in *IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 5865–5869.

[178] A. Tomasoni, M. Ferrari, D. Gatti, F. Osnato, and S. Bellini, "A low complexity turbo mmse receiver for w-lan mimo systems," in *Proc. IEEE Int'l Conf. Commun. (ICC)*, vol. 9, Jun. 2006, pp. 4119–4124.

[179] K. Li, C. Jeon, J. R. Cavallaro, and C. Studer, "Feedforward architectures for decentralized precoding in massive MU-MIMO systems," *arXiv:1804.10987 [cs.IT]*, 2018.

[180] P. Meinerzhagen, C. Roth, and A. Burg, "Towards generic low-power area-efficient standard cell based memory architectures," in *Proc. IEEE Int. Midwest Symp. Circuits and Syst. (MWSCAS)*, Aug. 2010, pp. 129–132.

[181] B. Gestner and D. V. Anderson, "Single Newton-Raphson iteration for integer-rounded divider for lattice reduction algorithms," in *Proc. IEEE Int. Midwest Symp. Circuits and Syst. (MWSCAS)*, Aug. 2008, pp. 966–969.

[182] P. Kyösti, J. Meinilä, L. Hentilä *et al.*, "WINNER II channel models. D1.1.2 V1.2," Tech. Rep. IST-4-027756 WINNER II, Sep. 2007.

[183] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[184] O. Castañeda, S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "1-bit massive MU-MIMO precoding in VLSI," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 4, pp. 508–522, Dec. 2017.

[185] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Apr. 2014.

[186] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, no. 6, pp. 1135–1151, Nov. 1981.

[187] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Aug. 1994.

[188] K. R. Kumar, G. Caire, and A. L. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4398–4418, Oct. 2009.

[189] C. D. Perels, "Frame-based MIMO-OFDM systems: impairment estimation and compensation," Ph.D. dissertation, ETH Zurich, Switzerland, 2007.

[190] H. Nishimori, *Statistical physics of spin glasses and information processing: an introduction*. Oxford University Press, 2001, no. 111.

[191] J. Barbier, M. Dia, N. Macris, and F. Krzakala, "The mutual information in random linear estimation," in *Proc. Allerton Conf. Commun., Contr., Comput.*, Sep. 2016, pp. 625–632.

[192] J. Barbier, N. Macris, M. Dia, and F. Krzakala, "Mutual information and optimality of approximate message-passing in random linear estimation," *arXiv:1701.05823 [cs.IT]*.

[193] C. Rush and R. Venkataramanan, "Finite-sample analysis of approximate message passing," in *Proc. IEEE Int'l Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 755–759.

[194] D. Guo, Y. Wu, S. Shamai, and S. Verdú, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.